ARO 26634.1-EL-CF

2

ADVANCED COMMUNICATION PROCESSING TECHNIQUES

by

Robert A. Scholtz

CSI-90-04-01

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	1990	Final 13 Feb 89 - 30 Jun 90

4. TITLE AND SUBTITLE
Advanced Communication Processing Techniques

5. FUNDING NUMBERS
DAAL03-89-G-0016

6. AUTHOR(S)
Robert A. Scholtz

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Univ of Southern California
Los Angeles, CA 90089-0272

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U. S. Army Research Office
P. O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
ARO 26634.1-EL-CF

11. SUPPLEMENTARY NOTES
The view, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

This document contains the proceedings of the workshop "Advanced Communication Processing Techniques," held May 14-17, 1989, near Ruidoso, New Mexico. Sponsored by the Army Research Office (under Contract DAAL03-89-G-0016) and organized by the Communication Sciences Institute of the University of Southern California, the workshop had as its objective "to determine those applications of intelligent/adaptive communication signal processing that have been realized and to define areas of future research."

We at the Communication Sciences Institute believe that there are two emerging areas which deserve considerably more study in the near future: (1) Modulation characterization, i.e., the automation of modulation format recognition so that a receiver can reliably demodulate a signal without using *a priori* information concerning the signal's structure, and (2)

Continued on back

14. SUBJECT TERMS
Workshop, Advanced Communication, Communication Signal Processing, Modulation Characterization, Adaptive Coding

15. NUMBER OF PAGES
476

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

ABSRACT CONTINUED:

the incorporation of adaptive coding into communication links and networks. (Encoders and decoders which can operate with a wide variety of codes exist, but the way to utilize and control them in links and networks is an issue.) To support these two new interest areas, one must have both a knowledge of (3) the kinds of channels and environments in which the systems must operate, and of (4) the latest adaptive equalization techniques which might be employed in these efforts.

The first two days of the workshop were occupied by four sessions, each session headed by a panel of acknowledged experts from industry and academia. Each panel was commissioned to survey past efforts in a particular problem area and to propose problems which need further study. The topics for these four sessions were (a) *military communication channels*, (b) *current issues in equalization*, (c) *adaptive coding for error correction*, and (d) *modulation characterization*. The fifth session of the workshop, held on the morning of the third day, was available for three purposes: (1) To extend, if desired, any discussions from the previous four sessions, (2) To summarize and critique the content of the workshop, propose other topics for future discussions, etc., and (3) To discuss perceived problems in carrying out future research, e.g., funding, communication of results, etc.

ADVANCED COMMUNICATION PROCESSING TECHNIQUES

The proceedings of a workshop sponsored by the

U.S. ARMY RESEARCH OFFICE
P.O. Box 12211
Research Triangle Park
North Carolina 27709

under Grant No. DAAL03-89-G-0016

and the

COMMUNICATION SCIENCES INSTITUTE
Department of Electrical Engineering
University of Southern California
Los Angeles, CA 90089-0272

May 14-17, 1989

TABLE OF CONTENTS

PREFACE

This document contains the proceedings of the workshop "Advanced Communication Processing Techniques," held May 14-17, 1989, near Ruidoso, New Mexico. Sponsored by the Army Research Office (under Contract DAAL03-89-G-0016) and organized by the Communication Sciences Institute of the University of Southern California, the workshop had as its objective "to determine those applications of intelligent/adaptive communication signal processing that have been realized and to define areas of future research."

We at the Communication Sciences Institute believe that there are two emerging areas which deserve considerably more study in the near future: (1) Modulation characterization, i.e., the automation of modulation format recognition so that a receiver can reliably demodulate a signal without using *a priori* information concerning the signal's structure, and (2) the incorporation of adaptive coding into communication links and networks. (Encoders and decoders which can operate with a wide variety of codes exist, but the way to utilize and control them in links and networks is an issue.) To support these two new interest areas, one must have both a knowledge of (3) the kinds of channels and environments in which the systems must operate, and of (4) the latest adaptive equalization techniques which might be employed in these efforts.

The first two days of the workshop were occupied by four sessions, each session headed by a panel of acknowledged experts from industry and academia. Each panel was commissioned to survey past efforts in a particular problem area and to propose problems which need further study. The topics for these four sessions were (a) *military communication channels*, (b) *current issues in equalization*, (c) *adaptive coding for error correction*, and (d) *modulation characterization*. The fifth session of the workshop, held on the morning of the third day, was available for three purposes: (1) To extend, if desired, any discussions from the previous four sessions, (2) To summarize and critique the content of the workshop, propose other topics for future discussions, etc., and (3) To discuss perceived problems in carrying out future research, e.g., funding, communication of results, etc.

All sessions were tape-recorded and transcribed in an effort to accurately preserve the sense of the discussions. Transcripts of presentations were edited for clarity and for inserting references to the speakers' slides, and the results were approved by the presenters. Hence the following proceedings are, for the most part, not formally prepared papers, but edited transcriptions.

Special thanks for the timely and high-quality production of these proceedings go to Cathy Cassells and Milly Montenegro, the workshop administrative assistants. Thanks for a job well done also go to Monisha Ghosh, David Rollins, and Michael Rude, who supervised the recording process, and edited the transcriptions of the sessions.

Dr. Robert A. Scholtz
Workshop Chairman

Modulation Characterization

Session Chairman: Charles Weber

Bart Rice
Automatic and Interactive Signal Classification

Mark Wickert
Modulation Characterization using Rate-Tone Generation Systems

Steve Stearns
Statistical Pattern Recognition versus Model-Based Approaches to Signal Classification

Edward Satorius
Application of Neural Networks to Signal Sorting

ROBERT SCHOLTZ: Welcome to the Workshop on Signal Processing Applications in Communications. This workshop is jointly sponsored by the USC Communication Sciences Institute and by the Army Research Office. Complaints or constructive criticisms, whatever ... please pass them on to me or to Dr. Sander (ARO) in the back – Bill, please wave your hand – and Jim Gault (ARO) – is Jim here? – Jim, over here, as well.

I'd like to say a few words about the process that we're undertaking here. We're going to transcribe the full proceedings of the workshop, that is, both the talks and the discussion. So we'd appreciate your help. When you have to ask questions, we have a mike here and we'll pass it to you. If possible remember to state who you are so that in the transcription we can even say who asked the question, unless you want to remain totally anonymous. If you really want to remain anonymous, write the question on a piece of paper and pass it to one of my helpers and they will get it up to the chairman of the panel who will ask the question.

All speakers, I would appreciate it if you would get copies of your viewgraphs to Milly. Milly, would you stand up? Milly is in the back of the room. She is the administrator of the whole workshop. If you have financial problems or whatever, talk to Milly. [LAUGHTER] She will also be corresponding with all the speakers, so you will get copies of our transcription for your editing before the publication of the proceedings. The whole process typically will probably take until about November, would you say Milly? December, that's usually when the proceedings will come out transcribed. edited and published. Every attendee will receive a copy in the mail.

I would also like to introduce my three helpers over here: Dave Rollins, Monisha Ghosh and Mike Rude. They're all Ph.D. students at USC; they're working in various aspects of some of the session topics. They may come up to you and ask you what your name is because they're trying to keep track of who is asking questions and that sort of thing, to make it easier for the transcriber to get that information into the record. So don't look funny at them if they look at your name tag or ask you who you are.

OK, with that I'd like to turn this first session over to Chuck Weber who'll introduce his panelists. We're not going to coffee break officially until all the formal presentations are over and these speakers get their five minutes' say about what they think the area is. After the coffee break, I really want to open the session to total discussion. So make some notes. If you have a penetrating question, write it down. If you have a little question, that's just, "Is that an x or a y, because I can't read it?" that's fine. Ask that during their presentation. But save the penetrating ones for the discussion time, and we're leaving well over an hour in each session for just interchange between panelists and attendees. Charles, you're on ....

CHARLES WEBER: Thank you, Robert, and good morning. I've told each of the panelists that I would give them 15 or 20 minutes each and I will take zero time for me. So the first thing we'll do is violate that. [LAUGHTER, PAUSE] This might even work!

In order to at least get us started, I have put together a viewgraph or two of what might be considered motherhood, addressing issues that may come up in the question session, addressing what the goal of modulation characterization is (in lots of noise!).

Modulation Characterization

Environment:

* **Signal Structure Known, Except for a Set of Parameters**
* **RCVR Noise - AWGN**
* **Possible Narrowband and/or Wideband Random Interferences**
* **Filtering Effect / Channel Distortion - Excess Modulation**

➠ ***Need for Adaptive Equalization***

* **Fading / Multipath**

➠ ***Need for Channel Characterization***

Modulation Characterization

Goal: **Exploit Profile of Selected Digital Signals to perform Det./ Est./ Classification**

Tools:
* **Likelihood Functionals & Their Tests**
* **Suboptimal Versions (GLF)**
* **Parameter Estimation**
* **Time/Frequency Correlation Tests**
  Ex. SPCR, Wigner-Ville Dist.
* **Ad-Hoc Schemes (e. g. Feature Extraction, Pattern Recognition)**
* **Neural Nets**

Criteria:
* **Prob. of Detection and False Alarm**
* **Variances of Parameter Estimators**
* **Prob. of Correct Classificsation and Rejection**

In the exposure I've had to modulation characterization, which is not extensive when compared to what some of you in the audience have had, I'm viewing it as developing methods and algorithms which exploit the profile of the selected digital – I guess it wouldn't have to be digital a signal, but that's how I'm looking at it – and to perform detection, estimation, and classification within a set of candidates. We could certainly use the maximum likelihood and generalized likelihood functionals, and suboptimal versions of those. In addition, there is parameter estimation of those parameters that we're not completely sure of in the waveform that we're characterizing. Then there is a variety of schemes, correlation tests, spectral correlation (which I have abbreviated SPCR) Wigner-Ville. There's a whole list of time-frequency correlation methods and then more ad hoc scheming, feature extraction, pattern recognition, and one I recently added, neural nets, because Ed Satorious is going to tell us a little bit about this before this morning is over, namely how neural nets will fit potentially in modulation characterization.

I've also listed a couple of criteria. The standard Neyman-Pearson approach for detection, variances for parameter estimation, and probability of direct classification and rejection, if it happens to be an environment which involves more than just Neyman-Pearson criterion. Hopefully there will be comments and additions to this before we're through.

The second and last viewgraph, which we can do in zero time, is to at least say something about the environment. Here we assume some candidate known signal is present except for possibly a set of parameters, receiver noise, a variety of interferences which might be narrowband and/or wideband, channel effects, channel distortions, and excess modulation.

One of the reasons for this viewgraph is to kind of give you a measure of how the workshop came about. We originally decided that it would be on modulation characterization, and came to the conclusion that we're really not going to be able to do four sessions on modulation characterization. So we wanted some sessions that would essentially be in areas that would complement modulation characterization. What came to our mind was adaptive equalization and channel characterization as areas that would be very closely connected to modulation characterization, and we believe that that's certainly the case.

OK, that's really as much as I wanted to say, just about, in terms of an introduction. The first speaker – the first real speaker – is Bart Rice from Lockheed, and he's going to tell us about the automatic and interactive – I put a few extra letters in there – signal classification. [LAUGHTER] Well, that's just a test for us to see whether we're really going to be able to pull this off .... Bart, can't you tell? [LAUGHTER] If it was anybody else, I wouldn't say ....

BART RICE: *Automatic and Interactive Signal Classification*

I'm going to talk principally about a project that had to do with automatic characterization of signals in voice channels. We've also done some work in the RF arena, where we've also characterized automatically the kinds of signals that can appear there. I think the relevant points can be made by talking about our VCPM, our "voice channel processing module" [VIEWGRAPH #2]. I guess the lesson on this picture [VIEWGRAPH #3] is that a system that did this kind of thing was actually built and worked, and performed

SIGNAL ANALYSIS

BART RICE

ELECTRONIC SYSTEMS ARCHITECTURE

VIEWGRAPH #1

VOICE CHANNEL

SIGNAL PROCESSING

VIEWGRAPH #2

all of the engineering functions as well as the algorithmic functions to recognize the different kinds of signals that could appear. We would take in up to 300 channels that were frequency division multiplexed (FDM). We had a transmultiplexer that would take an FDM baseband into TDM form. The "Continuous Channel Processor" was a box that performed 40-point transforms for the purpose of what we call coarse classification. We processed 300 channels in two Numerix 432 array processors, which are 30 Megaflops machines. So, that's a lot of data to process in a hurry, in a machine with modest capabilities.

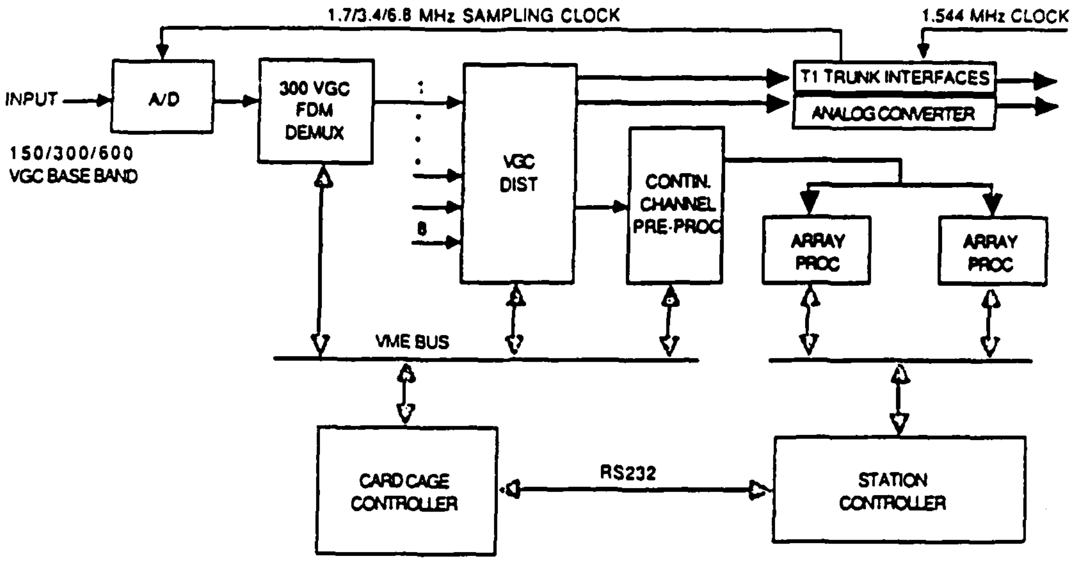
The first thing we did was to decide if the incoming signal was voice, because that would be the case most of the time. If it wasn't, then we went on to our "fine classifier." Data channels were dumped in the AP's and ping-ponged between a left memory to a right memory. While one memory was filling up the other memory was being processed in real time, and we had 150 channels going into each of the array processors.

The point of this [VIEWGRAPH #4] is that the method by which we derived the algorithms was heavily based on graphics and actually looking at various feature displays. That puts our techniques largely into the ad hoc category that Chuck was talking about. However, we also used some of the other, more systematic methods, and I'll touch on those. Basically, our approach was this: we identified a large number of features – and I'll tell you in a minute how we derived the features – we created the software that would create graphics and plot a variety of different displays; we put tape on the wall and we mounted the displays. We looked at these visual representations of the different modulation types, and we said "What is there about this particular modulation type that stands out?" We would look at those pictures and say something like, "Well, if it's an FSK and you put it through a frequency discriminator, you get a square wave that looks distinctive. How can we capture that in a feature that we could use in a clustering program to separate FSK from other modulation types?"

Another thing we did was to see how robust these features were once we identified them. We did that in two ways. We varied the signal-to-noise ratio and we also used three different channel filters. One filter was a pristine filter which passed everything with a flat passband and zero group delay. Another one was a filter that had given our demodulator developers a hard time on a previous project. It was their worst-case filter when trying to get the Godard algorithm to converge for V.29-modulated signals. It had an interesting and a very sharp group delay characteristic. Another one simply attenuated much of the upper-half band. It was, I think, representative of a lot of older telephone equipment that distorted the upper part of the passband pretty badly. So, we used those three filters and required that our features be robust, in the sense that their distributions didn't change much when the signals were passed through any one of those filters.

In our coarse classifier [VIEWGRAPH #5], we would perform interim decisions over a second. We would integrate those over about 10 seconds, although later on, in both our coarse and fine classifiers, some of the features would be actually computed over much shorter intervals. At the time we did this we weren't sure exactly how long feature computation would take. This chart [VIEWGRAPH #5] shows that the coarse classifier had interim decisions, integrated decisions, and there was a figure of merit associated

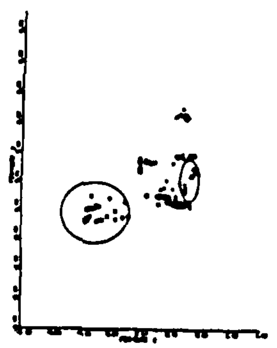
VCPM FUNCTIONAL ARCHITECTURE



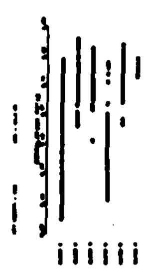
VIEWGRAPH #3

VOICE GRADE CHANNEL REALTIME SIGNAL CLASSIFICATION

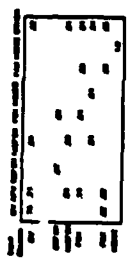
- FEATURE-BASED
- CLUSTER TO OBTAIN "INTERIM" DECISION
- RAPID "COARSE" CLASSIFICATION (VOICE VS. NON-VOICE)
- "FINE" CLASSIFIER DISCRIMINANTES BETWEEN SIGNAL CATEGORIES
  - CLEAR VOICE
  - FSK
  - PSK/QAM/PR
  - MULTI-CANAL SIGNALS
  - NOISE
  - OTHER



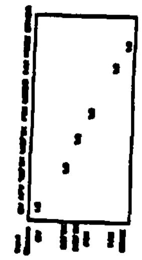
2-D SCATTER PLOT SHOWS HOW A PAIR OF FEATURES SEPARATES CLASSES



1-D DISTRIBUTION SHOWS HOW A SINGLE FEATURE SEPARATES CLASSES



INTERIM "CONFUSION MATRIX"

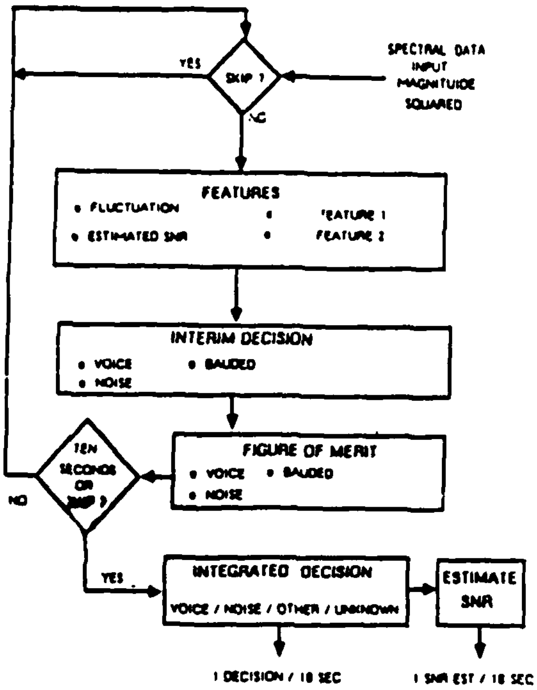


INTEGRATED "CONFUSION MATRIX"

- PRELIMINARY FEATURE SELECTION
  - IDENTIFY DISTINGUISHING CHARACTERISTICS OF EACH SIGNAL CLASS
  - QUANTIFY THESE CHARACTERISTICS AS A FEATURE
- FINAL FEATURE SELECTION
  - FEATURE DISTRIBUTIONS (ONE-DIMENSIONAL)
  - SCATTER PLOTS (TWO-DIMENSIONAL)
  - DISCRIMINANT ANALYSIS (MULTI-DIMENSIONAL)
  - ELIMINATE REDUNDANT FEATURES
  - ELIMINATE FEATURES WITH STATISTICS OVERLY SENSITIVE TO NOISE AND CHANNEL DISTORTION
- DECISION LOGIC
  - HIERARCHICAL
  - BAYESIAN
- IMPLEMENTATION
  - VAX/NUMERIX MARS 432 ARRAY PROCESSOR
  - CONVENTIONAL FORTRAN OR RULE-BASED

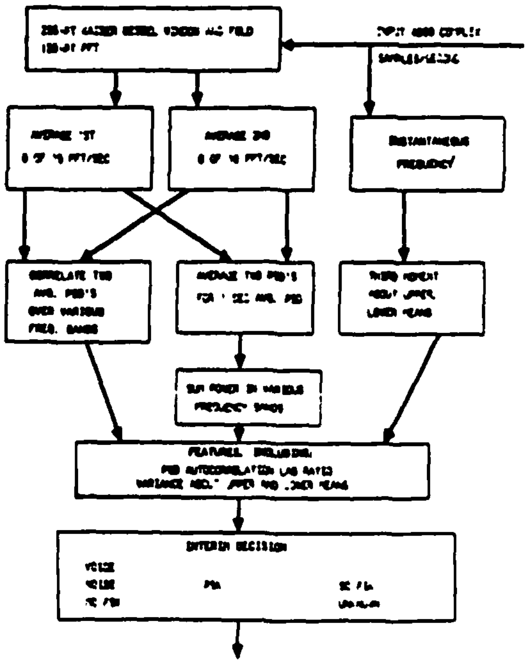
VIEWGRAPH #4

COARSE CLASSIFIER BLOCK DIAGRAM



VIEWGRAPH #5

FINE CLASSIFIER BLOCK DIAGRAM



VIEWGRAPH #6

with the quality of each decision.

The fine classifier [VIEWGRAPH #6] block diagram shows the flow of the algorithms, and it lists some of the features that were used. One of the problems we had was a system constraint – we were only sampling minimally. We had 4000 complex samples per second in a 4 kilohertz channel. Now, that means that some of the things you'd like to do, such as square the signal, look at the spectrum of the envelope, and compute various other non-linear things, cannot be done, because often when you do non-linear operations you expand the bandwidth. If you're minimally sampled that means that the output of the non-linear operation is aliased. So, we had to design our recognizers using only operations that didn't increase bandwidth. This was a case where we had to do some things that we knew were decidedly suboptimal, because of system constraints. This figure [VIEWGRAPH #6] also shows the various classes that we were classifying: voice, noise, frequency shift keyed, phase shift keyed (including QAMs and partial responses), unknown, and multichannel voice frequency telegraphy. It could be FSK, on-off keyed, or PSK in the individual channels.

In the fine classifier, we would perform the interim decisions and then we'd integrate [VIEWGRAPH #8]. Then we would perform a secondary integration, where we would establish what we call a "steady state." If we had gotten a certain number of integrated decisions in a row that were the same, we said, "OK, we're in a steady state in that modulation type," and, therefore, to declare a change – that is, declare that we'd gone off or declare a change to a different modulation – we had to have something extraordinary happen, like declaring a different class from the steady state many times in a row. That turned out to be a very useful idea when we were dealing with real-world signals and real-world channels.

Let me give you some examples of some of the features and the kinds of graphics that we used to try to detect these things. This figure [VIEWGRAPH #10] shows fluctuations in a voice signal. We took the envelope and plotted it. These are three solid lines that look like they go together. The middle line is a long term average, the lines above and below are thresholds above and below, and the wavy line is a short term average of the energy. We counted the number of threshold crossings in a certain interval. This technique came from an off-hand remark of John Treichler – that you could detect voice very quickly just from the fact that the energy fluctuated a lot. Once you think about it, it's obvious.

This is the same picture with a QPSK signal [LAUGHTER] and you see that there are no fluctuations at all, because it's relatively constant envelope. The short term averages were 10 millisecond averages, and the long term averages were 1 second averages.

We also used scatter plots [VIEWGRAPH #12] quite a lot, and what we found was that, almost always, two or three of the features were sufficient to characterize a given modulation type and separate it from the other classes. We could almost always capture that separation pictorially in one or two scatter plots. This scatter plot [VIEWGRAPH #12] shows the crossings versus SNR. It shows that you get a lot of crossings for the voice and very few for the bauded signals.

This is another scatter plot [VIEWGRAPH #13] for a different pair of features. It illustrates how the signals separate. This is a power spectrum of a simulation of an R.35 [VIEWGRAPH #14] multichannel telegra-

FEATURES

FEATURE-1—DELETED

FEATURE-2—DELETED (USED TO SEPARATE FAX)

FEATURE-3—PSD AUTOCORRELATION LAG RATIO

FEATURE-5—VARIANCE ABOUT UPPER AND LOWER FREQUENCIES

FEATURE-6—DELETED

FEATURE-7—LOW-BAND TIME SEQUENTIAL CORRELATION

FEATURE-8—ESTIMATED SNR

FEATURE-10—LOW FREQUENCY POWER FLUCTUATIONS

FEATURE-11—PERCENT HIGH FREQUENCY ENERGY

FEATURE-12—LOW BAND SNR

FEATURE-13—PSD LOW BAND AUTOCORRELATIOIN LAG RATIO

VIEWGRAPH #7

FINE CLASSIFIER

INTEGRATED DECISION AND FIGURE OF MERIT (FOM)

The FOM represents how close to a threshold a signal's features fall.

The integration length is variable. If the sum of FOMs and the number of consecutive occurrences of a decision over some k second interval are greater than set thresholds, then an integrated decision is reported. The next 9-k seconds of data are skipped.

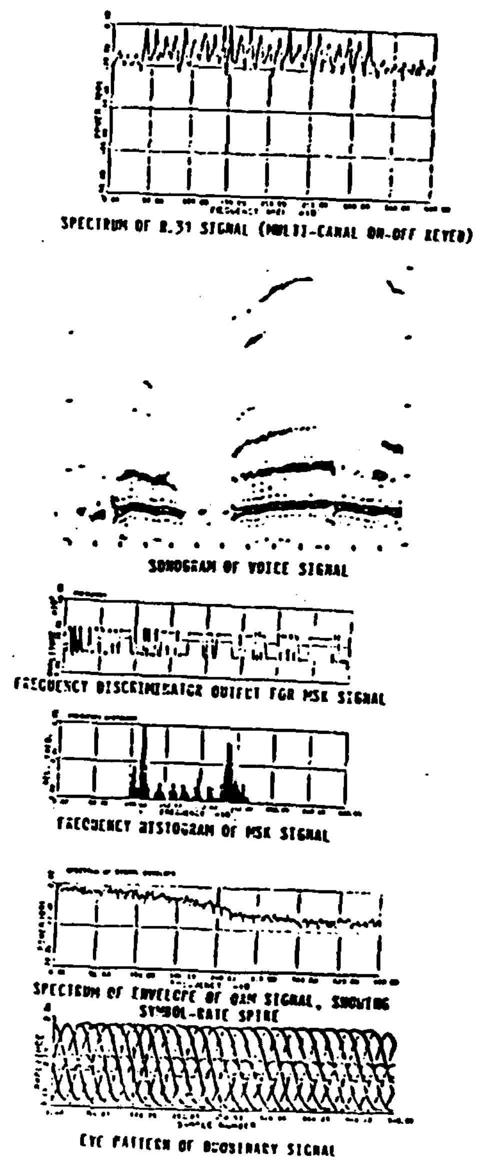
A final consideration is the SNR. If the SNR is below a "Poor Signal Quality" threshold, the integrated decision is "unknown" regardless of the FOM and consecutive number of Interim Decisions.

SECONDARY INTEGRATION

Secondary integration handles channel fades by using a transition matrix. Several matrices are available to assign to each channel.

VIEWGRAPH #8

WAVEFORM ANALYSIS



SPECTRUM OF 8.31 SIGNAL (MULTI-CANAL ON-OFF KEYED)

SONOGRAM OF VOICE SIGNAL

FREQUENCY DISCRIMINATOR OUTPUT FOR MSK SIGNAL

FREQUENCY HISTOGRAM OF MSK SIGNAL

SPECTRUM OF ENVELOPE OF QAM SIGNAL, SHOWING SYMBOL-RATE SPIKE

EYE PATTERN OF BIOSTINARY SIGNAL

- INTERACTIVE GRAPHICS
- IDENTIFICATION OF SPECIFIC MODULATION TYPE
- PARAMETER ESTIMATION (E.G., CARRIER, SYMBOL RATE)
- EQUALIZATION (BLIND AND *DECISION-DIRECTED*) AND DEMODULATION
- SIGNAL CONDITIONING
- HIGH RESOLUTION SPECTRAL ANALYSIS

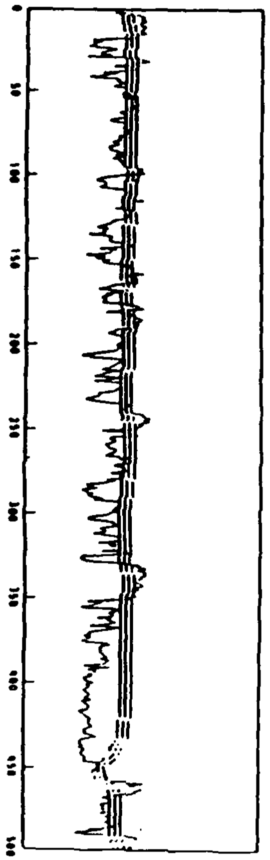
TOOLS:

- SOFTWARE
  - SASS-LMSC SIGNAL ANALYSIS, SIMULATION, AND MANIPULATION SYSTEM
  - COMSIM - LMSC COMMUNICATIONS AND SIMULATION PACKAGE
  - COSP - COHERENT SIGNAL PROCESSOR (RADAR)
  - COMMERCIAL PACKAGES

    IMSL SIG
    ISL IEEE
    DISSPLA
- HARDWARE
  - VAX 11/780
  - DATARAM
  - NUMERIX MARS 432 ARRAY PROCESSOR
  - LE CROY DIGITIZER
  - TEKTRONIX 4014 AND 9-VT100'S WITH RETRO-GRAPHICS

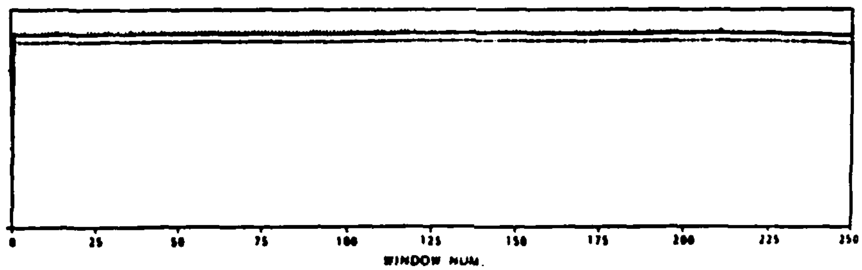
VIEWGRAPH #9

FLUCTUATIONS OF SPEECH SHORT-TERM ENERGY ABOUT LONG-TERM ENERGY



VIEWGRAPH #10

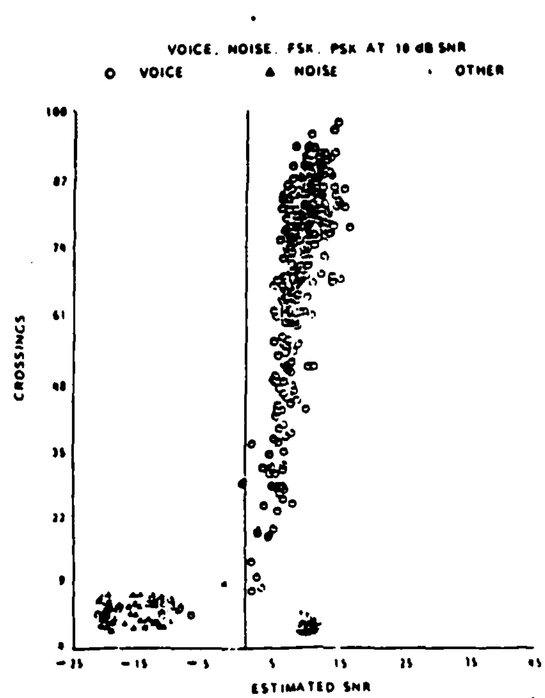
LONG, SHORT TERM POWER - QPSK



VIEWGRAPH #11

Lockheed
Missiles & Space Company, Inc

SCATTER PLOT



VIEWGRAPH #12

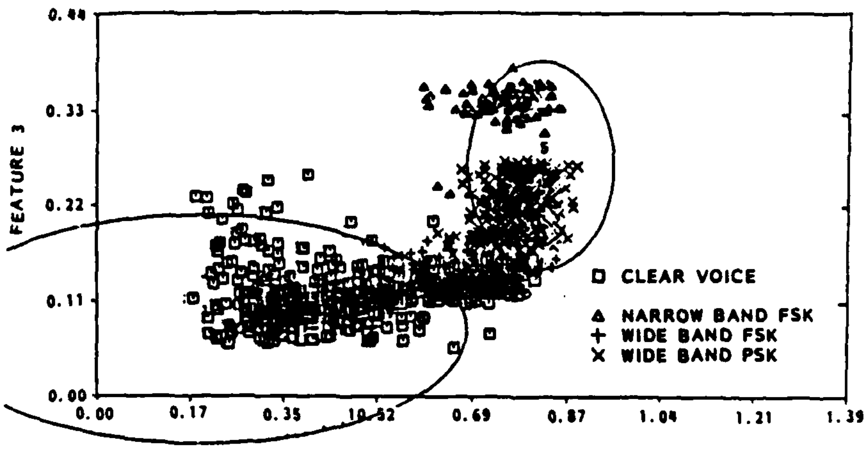
Lockheed
Missiles & Space Company, Inc

phy signal. What stands out to you about this? It's the peak/valley structure of the spectrum. How do you capture that? We observed that if we shift the picture so that the peaks line up over the valleys, then the correlation is low. Here's another one [VIEWGRAPH #15], for an R.31, in which the channels are on-off keyed, and this has some noise on it. The Cepstrum [VIEWGRAPH #16] also contains an obvious feature, but we did not use it because it was not robust and it required additional computation.

Here's a one dimensional plot [VIEWGRAPH #17] that we used. It shows the distributions of this feature worked for the different modulations. We took the spectrum and we correlated it against itself at an offset of about 30 hertz, which is a reasonable channel separation for these kinds of signals in this environment. And, you see that, for the multichannel signals, the correlation was much lower than it was for the single channel FSK and the PSK classes. Voice and analog fascimile were easily separated by other means. So, to separate the multichannel signal from the FSKs and the PSKs, this correlation technique did a very good job. By the way, this picture illustrates an important point. We would look at this picture and say, "We'll set a threshold at about 0.7. That will be our threshold for that feature, and it will separate the multichannel signals from the others." We might have been tempted to try something Bayesian and assume that the feature is Gaussian or has some known distribution with some mean. We might then look at the distance between the computed feature value and the mean. What we found out was that these distributions were very hard to characterize. The one here is bimodal. If we had run another case, the locations of the clusters might have been somewhat different, and the assumed distribution just wouldn't have been right. But, the threshold would have still been good. If you take a Bayesian approach, you get into likelihood ratio functions and probabilities of correct decisions based on the implicit assumption that the computed feature values are independent samples from the distribution. This picture [VIEWGRAPH #17] is based on an ensemble of signals. For a given MCFSK signal, the feature values would have come up consistently in one of the two clusters, and therefore the computed values were not independent. So, a given signal with given characteristics yields a distribution of the feature which may not be representative of the distribution as a whole. We'd have been fooling ourselves by computing a quality measure based on, say, successive feature values from the same signal and assuming that the feature values so obtained were representative for the whole ensemble. So, we didn't do it.

Of course, sometimes you can select features from the spectrum of the signal itself. There's an FSK with integer modulation index [VIEWGRAPH #18] and, of course, there are very distinctive peaks right at the two keying frequencies. But, when you look at the spectrum of an MSK [VIEWGRAPH #19], you just see a lump. It just looks spectrally like any other PSK or QAM signal. But, if you look at the spectrum of the square you see two peaks, because the square is an FSK with modulation index 1. Unfortunately, we couldn't compute the square. We were constrained by the minimal 4 kilohertz sampling rate. Note, by the way, that the scales are different on the plots of the spectrum and the spectrum of the square. The spectrum of the square is compressed by a factor of 2 relative to the spectrum. This is a useful display, because halfway between

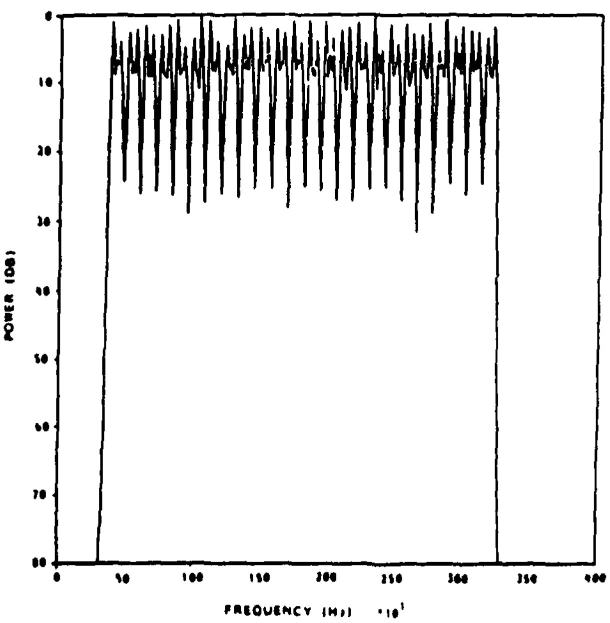
SCATTER PLOTS AS A TOOL IN SIGNAL ANALYSIS



VIEWGRAPH #13

Lockheed Missiles & Space Company Inc

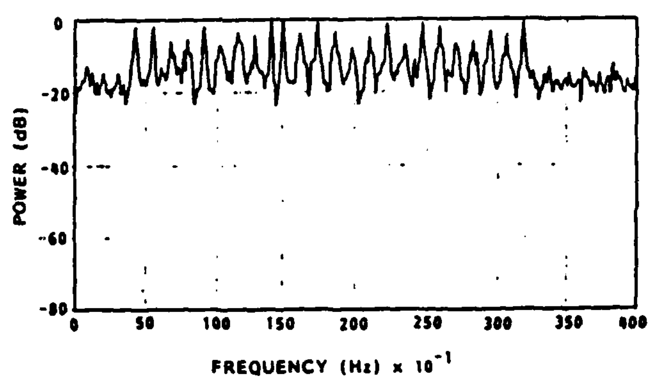
R.35 MCFSK POWER SPECTRUM



VIEWGRAPH #14

Lockheed Missiles & Space Company Inc

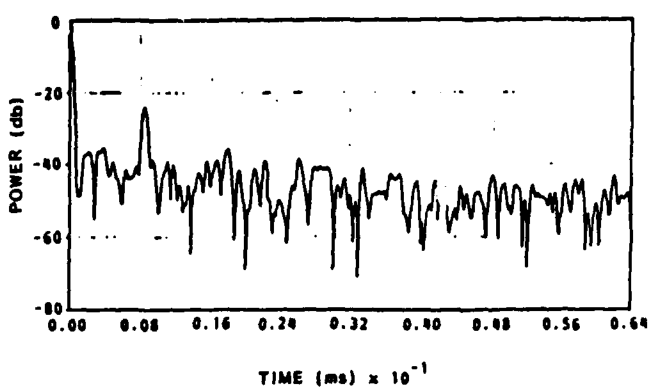
PSD AND CEPSTRUM OF R.31 MCVFT OOK



VIEWGRAPH #15

Lockheed Missiles & Space Company, Inc

CEPSTRUM OF R.31 MCVFT OOK

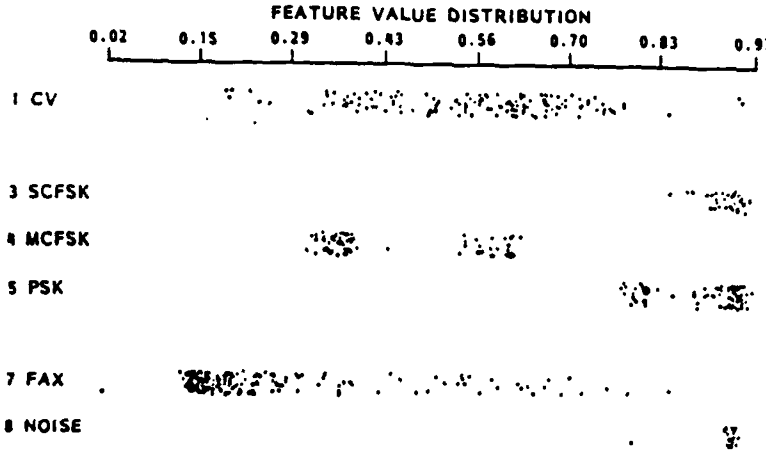


VIEWGRAPH #16

Lockheed Missiles & Space Company, Inc

PSD AUTOCORRELATION LAG RATIO

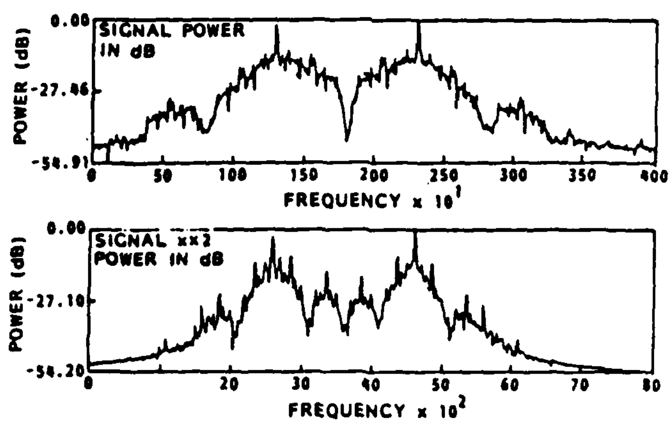
SNR = 40.0 dB



VIEWGRAPH #17

Lockheed
Missiles & Space Company Inc

PSDs OF POWER OF SLOW FSK



VIEWGRAPH #18

Lockheed
Missiles & Space Company Inc

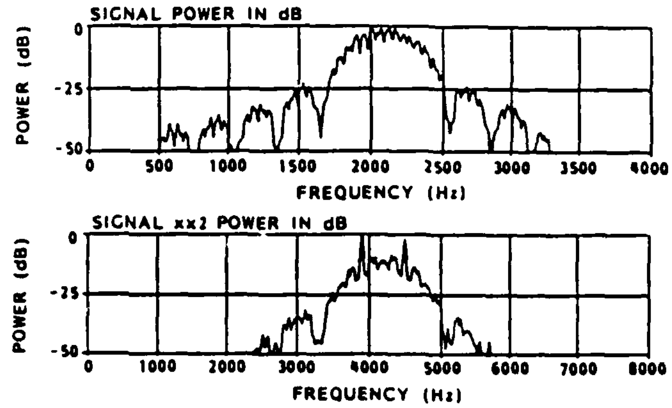
the two peaks is twice the carrier, which now lines up right at the center of the spectrum above. But, what really stands out with MSK or with any FSK signal is the square wave that comes out of a frequency discriminator [VIEWGRAPH #20]. If you take some noisier ones you can clean those up a little bit with a median filter [VIEWGRAPH #21]. Here's an example of a noisy one where you can see the 0's and 1's, but when you put it through a median filter it is much cleaner. The signal was oversampled, and the median filter contained 21 points. If you turn either of these last charts sideways and sum "down the columns," you get what's called a "frequency histogram." [VIEWGRAPH #22] Now, that's a very distinctive picture, this two-level frequency histogram. To capture this bimodality of the frequency histogram in a feature, we computed the mean, and then we computed what we called an "upper mean" and a "lower mean." The upper mean was the mean of the values that were above the mean, and the the lower mean was the mean of all the values that were below the mean. Then we computed the variance about the upper mean and the lower mean, and we added those together to obtain the feature.

The one-dimensional distribution [VIEWGRAPH #23] shows that, for FSKs, the variance was much less than for the other modulation types. And, again, every one of these features that we chose was robust. The feature distribution varied relatively little, the face of considerable channel distortion and noise.

In the development of our automatic signal classification algorithms, we also did a lot of signal analysis. We developed a signal analysis software package in which we were not constrained by not being able to square the signal or look at the spectrum of the envelope. I could point out a number of other features that one can exploit in the signal analysis mode. Quite a few of those are described in the attached paper, "Automatic and Interactive Signal Analysis." For example, in a bauded signal you see in the spectrum of the envelope, or in the spectrum of the output of a delay-and-multiply process, a line at the symbol rate. In the spectrum of the envelope in a voice channel of a V.29 quadrature amplitude modulated (QAM) signal, you see a baud rate spike come up at 2400 hertz [VIEWGRAPH #25]. The presence of a baud rate spike is one indicator that the signal is a PSK or QAM. We also look at the spectral shape and bandwidth and the relationship of that baud rate spike to the bandwidth. Those are other features for deciding that you have a QAM or a PSK signal. If you try to demodulate it, you don't obtain a recognizable "constellation." But, *if you put the signal* through a Godard equalizer, it starts to converge and you start to see the pattern emerge. The carrier must be "despun" also. In this picture [VIEWGRAPH #27], you can just barely see a constellation leaking through; but, if you run an automatic clustering algorithm [VIEWGRAPH #28], it'll converge to 16 clusters. You could then identify the specific modulation type on the basis of that convergence. This is a 4×4 QAM.

As mentioned before, you use not only a blind equalizer but also to despin the carrier. In blind equalizer mode we used a carrier tracking technique that I guess one of our fellow engineers rediscovered, but which Jim Mulligan developed several years ago. It's equivalent to one of the cases of the Goursat-Benveniste blind equalizer. Once the blind equalizer has converged and the constellation has been recognized, you can proceed into "decision directed mode." We have a

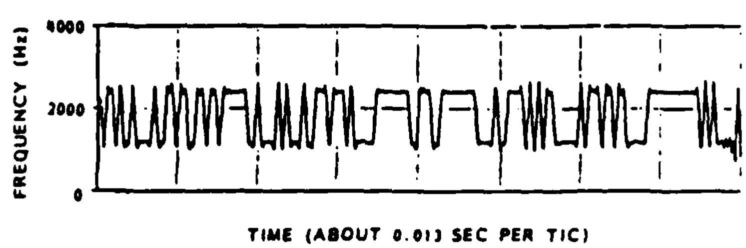
SPECTRA OF POWERS OF MSK



VIEWGRAPH #19

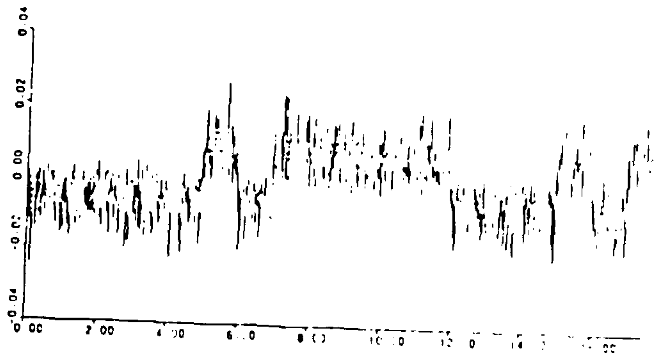
Lockheed Missiles & Space Company Inc

FREQUENCY DISCRIMINATOR OUTPUT OF MSK

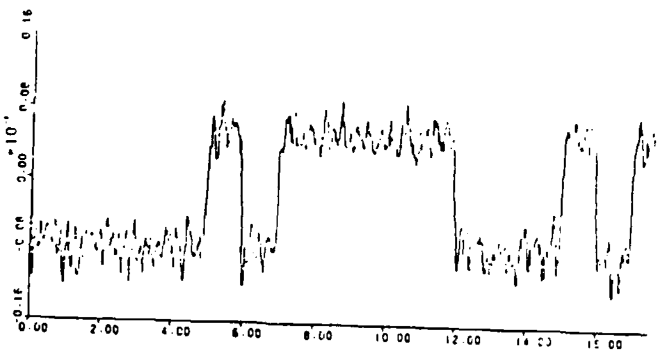


VIEWGRAPH #20

Lockheed Missiles & Space Company Inc

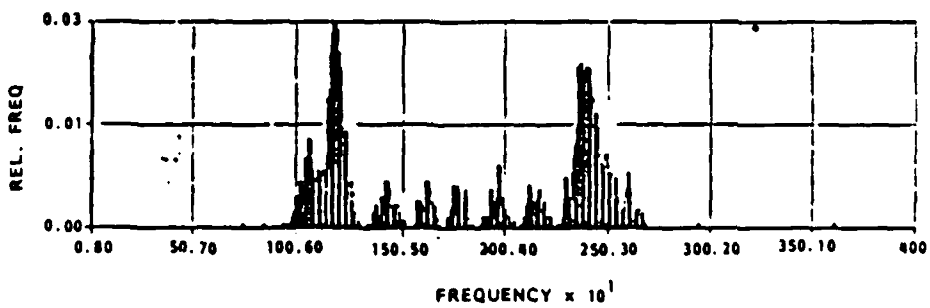


DEMODULATED FSK AFTER THE MEDIAN
FILTER IS APPLIED



VIEWGRAPH #21

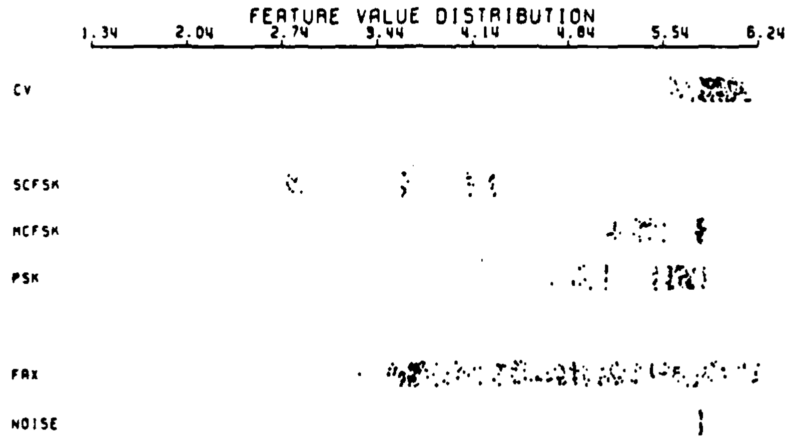
FREQUENCY HISTOGRAM OF MSK



VIEWGRAPH #22

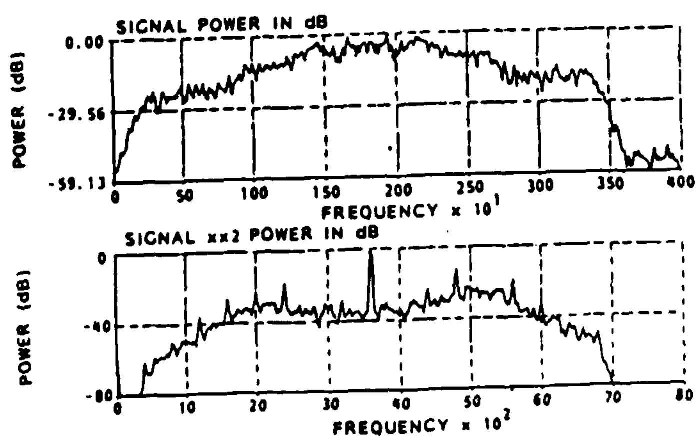
Lockheed
Missiles & Space Company Inc

VARIANCE ABOUT UPPER, LOWER FREQUENCIES



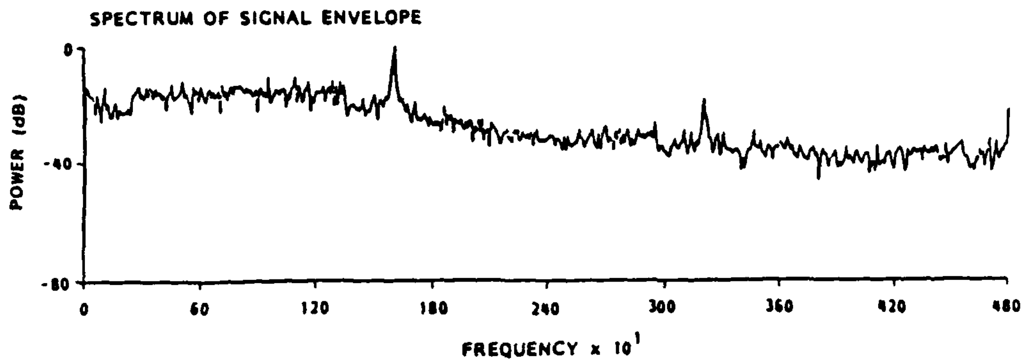
VIEWGRAPH #23

PSDs OF POWERS OF BPSK



VIEWGRAPH #24

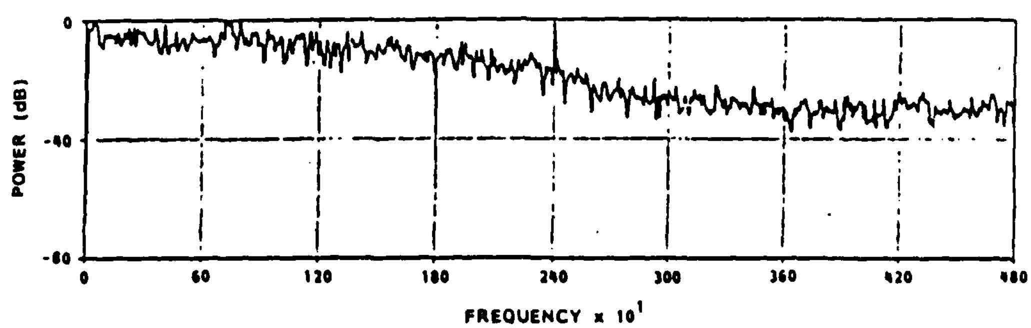
1600Hz SYMBOL RATE SUGGESTIVE OF V.27



VIEWGRAPH #25

Lockheed Missiles & Space Company Inc

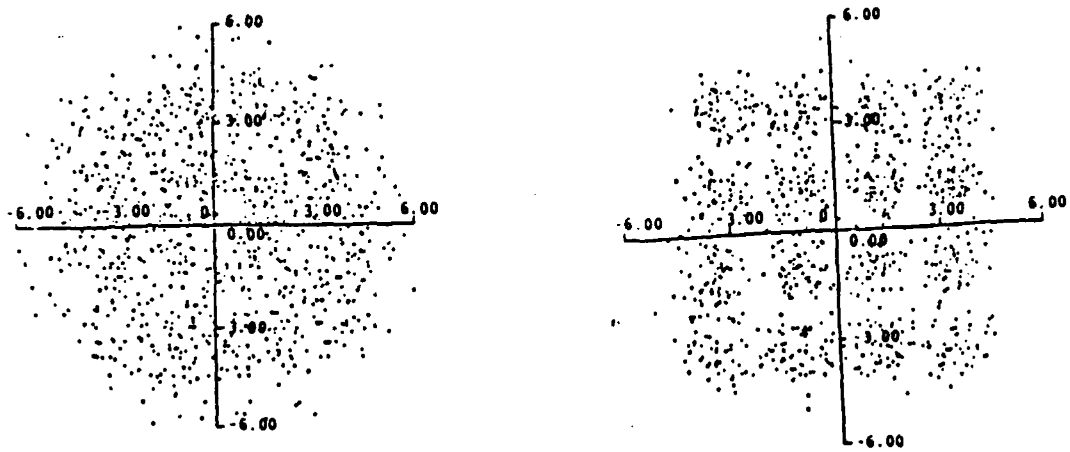
SPECTRUM OF SIGNAL ENVELOPE V.29



VIEWGRAPH #26

Lockheed Missiles & Space Company Inc

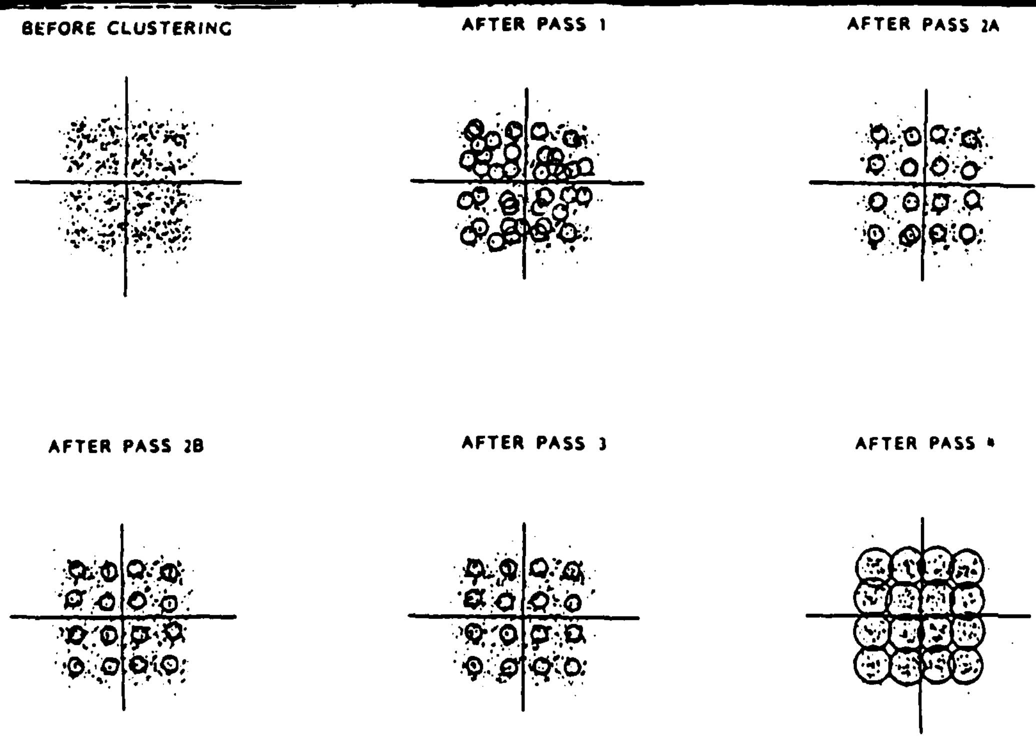
SNOWSTORM AND CONVERGED CASE



VIEWGRAPH #27

Lockheed
Missiles & Space Company Inc

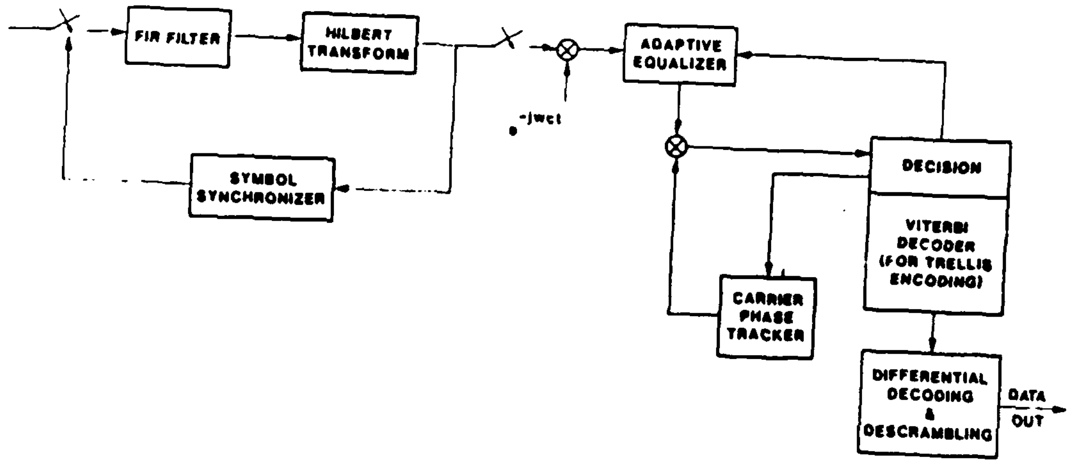
PICTORIAL EXAMPLE SHOWING CLUSTERING ALGORITHM PROCESSING STEPS



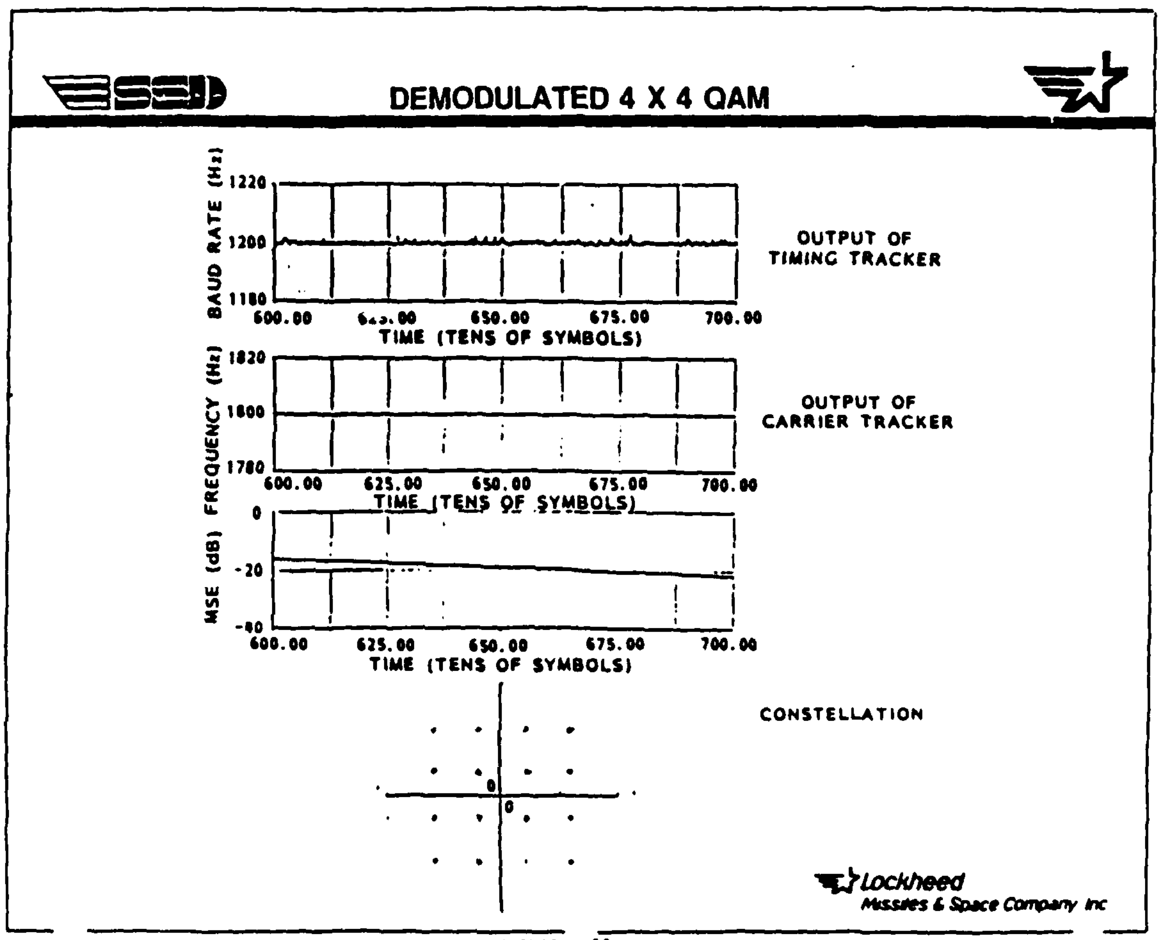
VIEWGRAPH #28

Lockheed
Missiles & Space Company Inc

GENERALIZED QAM DEMODULATOR



VIEWGRAPH #29



VIEWGRAPH #30

PARTIAL RESPONSE SIGNALS

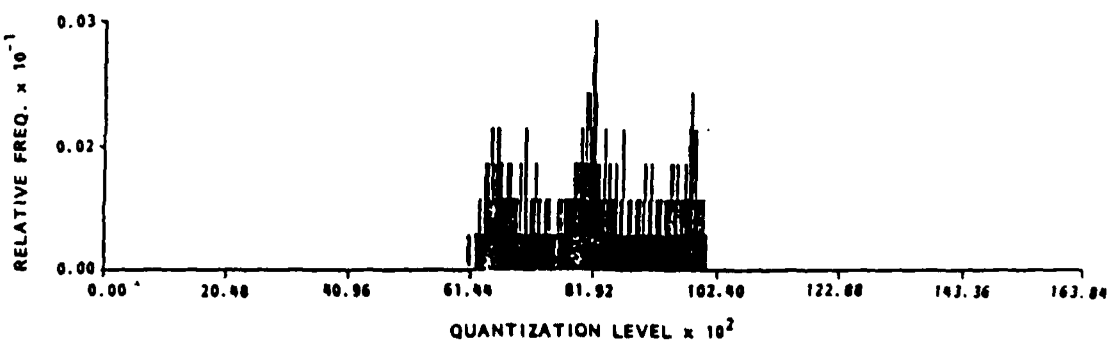
THESE MODULATION TYPES POSE A DIFFICULT CHALLENGE FOR THE SIGNAL ANALYST. THE PRINCIPAL SIGNALS IN THIS CLASS ARE TERMED "DUOBINARY" (CLASS 1), "MODIFIED DUOBINARY" (CLASS 4), AND "QUADRATURE PARTIAL RESPONSE" (QPR). OTHERS INCLUDE CLASSES 2 AND 3 AND "OFFSET QUADRATURE PARTIAL RESPONSE" (OQPR)

WHEREAS DESIGNERS OF COMMUNICATIONS SYSTEMS EMPLOYING QAM MODULATIONS ATTEMPT TO MINIMIZE INTERSYMBOL INTERFERENCE BY "PULSE SHAPING," PARTIAL RESPONSE MODULATIONS REQUIRE THE DELIBERATE IMPOSITION OF "CONTROLLED" INTERSYMBOL INTERFERENCE, WHICH MAY BE UNDONE AT THE RECEIVER. THE RESULT IS VERY SMOOTH SYMBOL TRANSITIONS AND REMARKABLY SPECTRAL EFFICIENCY - 2 SYMBOLS PER SECOND PER HERTZ OF BANDWIDTH - THE THEORETICAL LIMIT

VIEWGRAPH #31

Lockheed
Missiles & Space Company Inc

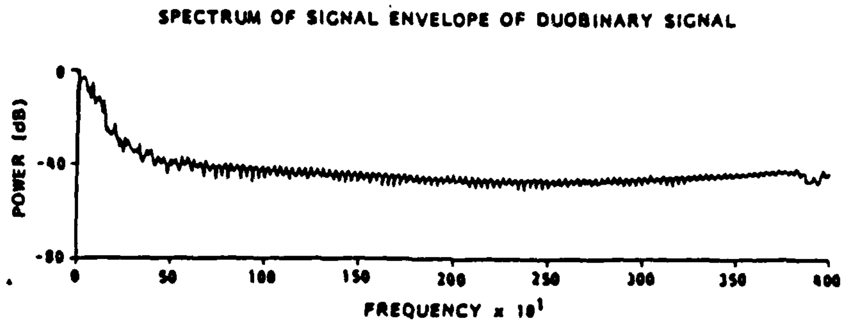
HISTOGRAM OF QUANTIZED SIGNAL BASEBAND DUOBINARY



VIEWGRAPH #32

Lockheed
Missiles & Space Company Inc

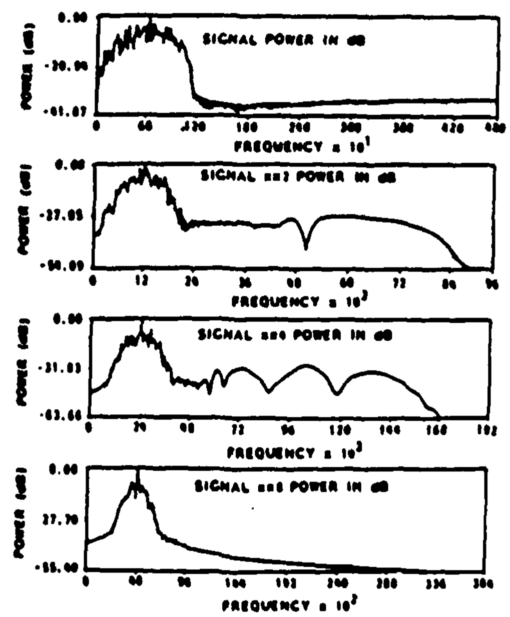
SPECTRUM OF SIGNAL ENVELOPE OF DUOBINARY SIGNAL



VIEWGRAPH #33

Lockheed
Missiles & Space Company Inc

PSDs OF POWERS OF A MODIFIED DUOBINARY SIGNAL



VIEWGRAPH #34

Lockheed
Missiles & Space Company Inc

PROPERTIES OF SIGNALS IN 4 kHz VGCs	VOICE	MCVFT	FSK	MSK	BPSK	PSK/VSB	QPSK	DQPSK	OQPSK	8-PSK	QAM	M-DUO/SSB	DUO/QPR	A-FAX
envelope fluctuations	yes													yes
spectral shape	spikey	peak-valley			2 peaks							log sin	log cos	
keying spike(s) in PSD of envelope/delay & mult		many	yes	yes	yes	weak (yes at baseband)	yes	yes	2xBW	yes	yes	NO	NO	
histogram	narrow		u-shaped									triangular at baseband		
spike(s) in PSD	yes		yes			Bell 203						often		
spike(s) in PSD of square				2	carrier	carrier								
spike(s) in PSD of 4th power				possibly			carrier	2	carrier			possibly 4xcarrier+BR		
spikes in PSD of 8th power										weak carrier	sometimes but weak			
instantaneous frequency			2-level			shows phase shifts								often constant

VIEWGRAPH #35

Properties of Different Signal Types

software QAM demodulator [VIEWGRAPH #29] that then converges further. You can see the trackers converge and go into tight constellations [VIEWGRAPH #30].

There are quite a few other things we could talk about here You have to do some different things for duo-binary signals. [VIEWGRAPHS #31, 32, 33] You don't see the spike in the spectrum of the envelope but you see something in the 4th power. [VIEWGRAPH #34]

Let me just close with a chart here that summarizes. This chart [VIEWGRAPH #35] shows a number of properties of signals that can be exploited for purposes of recognition: for voice, voice frequency telegraphy, frequency shift keyed, MSK. Besides a spike in the PSD of the 8th power of a quadrant-symmetric QAM, you'll also see a spike in the 4th power. You have to compute a long FFT to pull it out.

You can see that our approach to signal classification was largely ad hoc, it was heavily graphics-based, and it worked in a real-time system. There are several things that have been done since – neural nets, some of the work that Jerry Mendel has been doing at USC on higher order cumulants as a source for signal features – that one can use for clustering. This work represents some fertile ground for further research.

MARK WICKERT: *Modulation Characterization using Rate-Tone Generation Systems*

I want to begin with an outline of some of the things that I will talk about. [VIEWGRAPH #1] This is going to be considerably more narrower in scope than what the first talk was.

First I just want to go through, as an overview of delay and multiply receivers which, I guess, is something that's not really new, but just some of the basics and some of the things that you can do for getting rate tones out of digitally modulated signals. [VIEWGRAPH #2] So we'll look at the performance characterization, we'll look at the bandwidth and delay optimization of the basic receiver, we'll look at way of making the receiver not perform as well by using bandlimiting pulse shapes. I've got one slide on just looking at multipath channel performance. Then we'll go into more complicated receiver structures which use power series models for the nonlinearity. Next we'll look at the performance of these receivers with very narrowband signals, we could also consider an optimum structure. A lot of things have been done here, but just the highlights are presented here. The last topic considers a more exotic transmission scheme or communicator scheme, that will defeat even a 4th power nonlinearity if you combine both pulse shaping and amplitude weighting. So basically what's going on all the way through this is looking at it from two points of view. One is a person trying to gather information, and the second point of view is the communicator trying to make it more difficult for that person to get his information about your signal.

Here we have the abstract that I originally submitted for the workshop. [VIEWGRAPH #3] It's just telling you that rate tone generation of digitally modulated signals take advantage of the cyclostationarity properties of the received signal to generate a rate tone. These receiving systems work well for this purpose, but if you design your signals properly you can make these systems not work well at all so that you can look at it from both points of view. Some of the important things are considering the modulation type, types of nonlinear operators, pulse shaping, and using

Modulation Characterization Using Rate Tone Generation Systems

Mark A. Wickert
Electrical and Computer Engineering Department
University of Colorado at Colorado Springs
Colorado Springs, Colorado 80933-7150

Electrical and Computer Engineering Department
University of Colorado at Colorado Springs

OUTLINE

- Delay and Multiply Receiver
  - Performance Characterization
  - Bandwidth and Delay Optimization
  - Bandlimiting Pulse Shapes
  - Multipath Channels
- Power Series Nonlinearity
  - Performance Characterization
  - Narrow Bandwidth Signal Detector Law
- Combined Pulse Shaping and Amplitude Weighting
  - Gaussian Symbols
  - Quantized Symbols
  - Optimum Four-Level Signal

VIEWGRAPH #1

amplitude weightings. The last item, something I'm not really going to talk about, but something that does come up, is the type of spectral analysis (estimation) you would use to try to extract the features.

This is just to define the delay and multiply receiver. [VIEWGRAPH #4] The basic structure, this is a lowpass model – it could be bandpass, this just happens to be the way this figure was drawn – where the received signal is, in this case, a binary phase shift keyed signal in additive white Gaussian noise. You have two main parameters that you can adjust at the receiver: the front-end prefiltering bandwidth and the time delay of the delay multiply itself. The basic spectrum that you observe at the output $y(t)$ consists of a desired rate line component plus noise terms. In this case c_1 is the Fourier coefficient of *the fundamental rate tone*. There are several spectral components that are related to the noise terms that you get when you go through the delay and multiply, the self-noise or signal $\times$ signal (self-noise), signal $\times$ noise, and then noise $\times$ noise type terms, which all contribute to the background noise that you have to deal with when you're trying to resolve the spectral line. The equation that gives you the fundamental rate tone is this convolution type integral in the frequency domain. The reason why I'm displaying this is that later on this relationship plays a very important role in being able to make that rate tone go away. In other words, this is the transform of the basic pulse shape that comes through the prefilter, and then this is it shifted over by whatever harmonic at the rate tone that you're looking at. We'll see later that if you just make these bandlimited pulses and have it bandlimited to within one rate tone shift, you get no overlaps. You can then make the rate tone go away.

This is just a look at what the output is for a high signal-to-noise ratio case [VIEWGRAPH #5], just so it looks nicer. This is the spectrum coming out of the delay multiply, and this is the fundamental rate tone, the one that you're probably most interested in. This would be the third harmonic, this is for a BPSK signal. The background at this point is all self-noise because here we have a very high signal-to-noise ratio. This was just an average periodogram spectral estimate.

For most of the analysis that we were doing, we wanted to be able to ignore some of the noise components because analytically they're extremely difficult to calculate when you get to higher order nonlinearities. So in an attempt to justify some of that, we have – this plot shows the different contributions of the different noise terms that showed up in the denominator on the previous slide. If you're operating with a low $\frac{E_b}{N_0}$ at the front end of the system, say -10 dB or lower, you can neglect the self-noise and you can also neglect the signal $\times$ noise. So most of the time were were operating in that regime. In developing our performance measure signal-to-noise ratio, a model that just has that one noise term in it is used, thus things are a lot easier. It's not trivial but it's a lot easier.

I guess another thing I didn't mention is that signal-to-noise ratios may not seem like a valid measure, but some of the analyses that we had done involved looking at receiver operating characteristics. The statistics are approximately Gaussian for large time bandwidth products, so that that was a reasonable measure of performance. We weren't working with probabilities of detection and false alarm except at the very beginning.

OK, these are just some of the basic curves that you'd get when trying to optimize the receiver. [VIEWGRAPH #6] The receiver has

INTRODUCTION

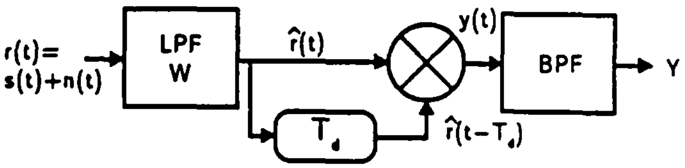
- Rate tone generation systems (circuits) take advantage of the cyclostationarity of digitally modulated carriers to facilitate detection and estimation of the symbol rate parameter. These systems provide good performance for signal interception applications, however, the communicator can significantly reduce the probability of being intercepted by such a system with careful signal design.
- The following issues are important in the study of these systems:
  - Signal Modulation Type
  - Types of Nonlinear Operators
  - Pulse Shaping
  - Amplitude Weightings
  - Spectrum Analysis Techniques

VIEWGRAPH #2

DELAY AND MULTIPLY RECEIVER

$s(t) = A\sum_k a_k p(t - kT_s)$

$n(t)$ = AWGN with PSD $N_o/2$



- The Detectability Performance Measure is the SNR at the Output of the Second BPF (spectral analysis operator)

$$\mathrm{SNR}_o = \frac{|c_1|^2}{B\left[\left(S_{s\times s}\left(\frac{1}{T_s}\right) + S_{s\times n}\left(\frac{1}{T_s}\right) + S_{n\times n}\left(\frac{1}{T_s}\right)\right)\right]}$$

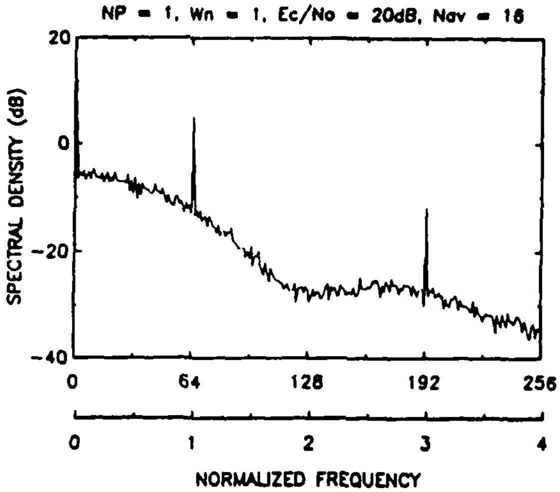
where

$$\begin{aligned}
|c_1| &= \text{Rate-Line Amplitude} \\
&= \frac{A^2}{T_s}\left|\int_{-\infty}^{\infty} P(f)P\left(\frac{n}{T_s} - f\right) e^{-j2\pi f T_d}\, df\right| \\
S_{s\times s}(f) &= \text{Signal times Signal (Self-Noise) PSD} \\
S_{s\times n}(f) &= \text{Signal times Noise PSD} \\
S_{n\times n}(f) &= \text{Noise times Noise PSD}
\end{aligned}$$

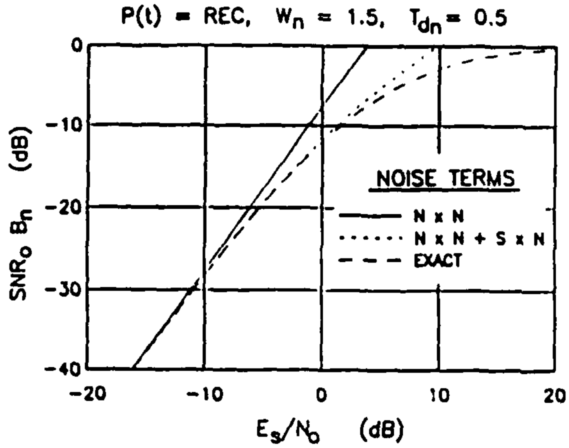
VIEWGRAPH #3

DELAY AND MULTIPLY RECEIVER CONT.

- The Power Spectrum at e(t) using Averaged Periodogram Spectral Estimation. Note that the background noise is primarily *self-noise* at high E_s/N_o levels.



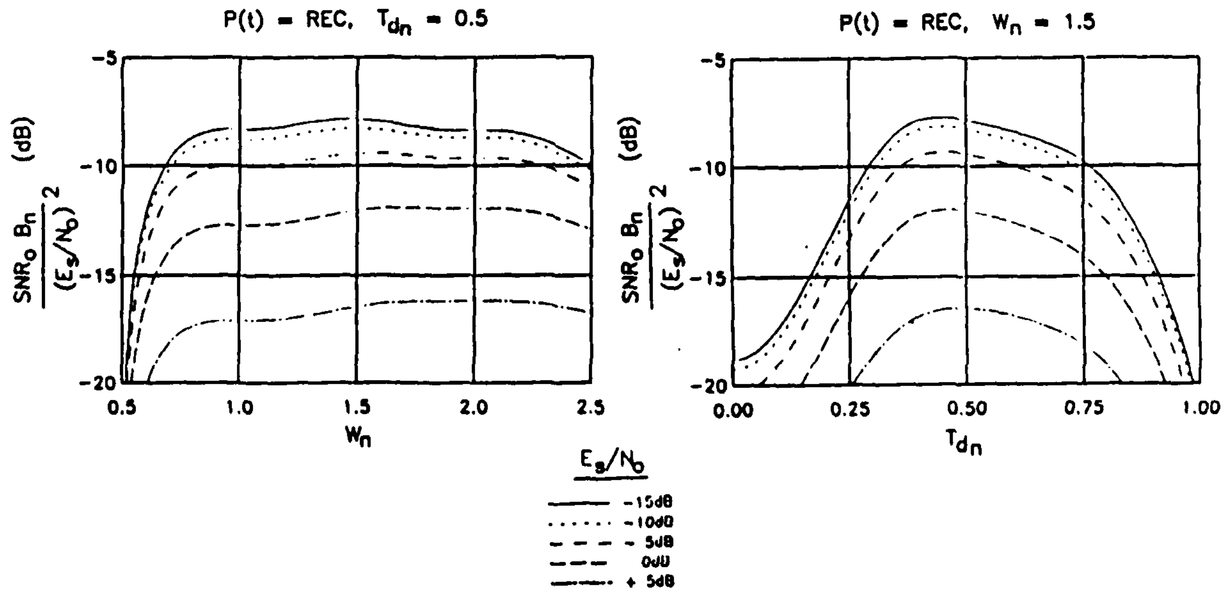
- For Intercept Receiver Applications where low E_s/N_o values are expected it is reasonable to neglect the *self-noise* and very often also the signal-times-noise term



VIEWGRAPH #4

DELAY AND MULTIPLY RECEIVER: *PERFORMANCE*

- Performance with BPSK: Optimization of the Receiver Parameters W and T_d



VIEWGRAPH #5

to look at a particular spot in the frequency domain to pick up your modulated carrier, and then it has to set the parameters of the bandwidth of the front end filter, then set the delay time. This is just showing you what the optimum bandwidth normalized to the symbol rate is. It's about 1.5 and it's fairly broad, and the optimum time delay is about 0.5. This is for rectangular pulse shapes, the standard types used for BPSK and QPSK. So this just gives an idea where an operating point might be for those types of modulation schemes.

Maybe I should explain this normalization because that's been a question before. The $(\frac{E_s}{N_0})^2$ is just thrown in the denominator there so the graphs can all appear stacked within reasonable spacings of each other. Because it's a squaring or a quadratic type operation that you're going through, when the signal-to-noise ratio drops you get the square of the signal-to-noise ratio dropping the SNR at the output. So that just keeps the curves from getting too widely spaced. But they're basically all the same shape, even when you drop down in $\frac{E_s}{N_0}$.

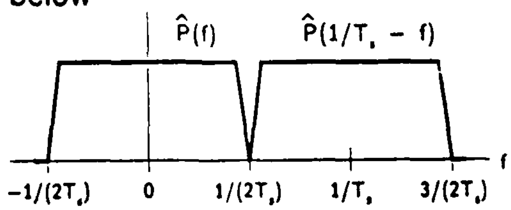
OK, the next thing is just look at what happens if you do some pulse shaping. [VIEWGRAPH #7] This is looking at a plot of basically the integrand of that expression for the rate tone. [VIEWGRAPH #4] This is just some hypothetical frequency spectra for the pulse shapes showing non-overlapping pulse spectra. Something that we could easily model just for analytical purposes was a Nyquist type pulse shape, one that has a sinusoidal roll-off characteristic. It has a single parameter, α, for its excess bandwidth and if you let α go to zero then you get a rectangular spectrum which is limited to half the symbol rate. You also have a $\frac{\sin x}{x}$ shape for the pulse shape which is not something you'd ever realize, but just have a limiting case where you could get the rate line to go to zero. Then you could back off on α or make it closer to 1 and you'd have a more gradual bandwidth roll-off. So that was hypothesized as a means of just looking at the performance of the delay and multiply with that type of pulse shape, or at least something that gives you a variable parameter allowing one to go from wideband to narrowband. There are lots of results compiled.

This is just one example in terms of how your receiver would have to operate if this pulse shaping was occurring. [VIEWGRAPH #8] This is just looking at as you change α. Obviously when α goes to zero you get nothing because the overlap goes to zero. But the more important thing that you can observe from this is related to the bandwidth parameter on the receiver front end. Recall that 1.5 was about in the middle of the optimum region for the rectangular pulse. Well now that you've put some pulse shaping on your transmitted signal, if you go with 1.5 and you gradually decrease α, you can see that your performance is going down rather rapidly. Whereas if you'd reoptimize the receiver and have a narrow bandwidth for that particular section that you're searching over, say to .6, you can sort of flatten it out and maintain slightly better performance. I guess what this is saying is that if you go to a non-rectangular pulse shape you have more sensitivity in the receiver parameters, so if you're trying to optimize you'd have to be more careful and maybe do finer levels of searching or something, just because of the fact that it's more sensitive and you're going to loose more. You can't have as broad a region to stay optimum over.

The power series nonlinearity is a means of trying to get back things that you've lost

DELAY AND MULTIPLY: *PULSE SHAPING*

- The Rate Tone Strength can be Reduced by Minimizing the Pulse Spectrum Overlap as Shown below

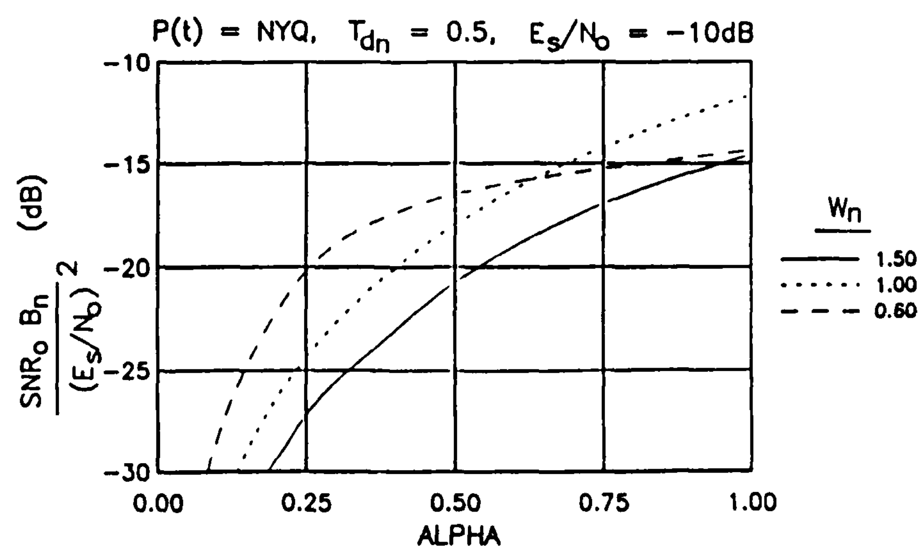


- A Nyquist Pulse, with Excess Bandwidth Parameter α, can be used to Create a Transmitted Signal with Bandlimited Spectrum having a Sinusoidal Roll-off Characteristic
- When $\alpha = 0$ the Spectrum is a Rectangle with Lowpass Bandwidth 1/(2Ts) and the Rate Tone Amplitude Becomes Zero

VIEWGRAPH #6

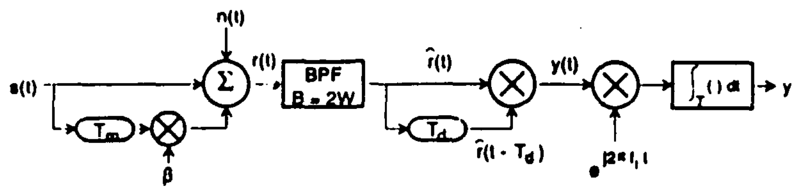
DELAY AND MULTIPLY: *PULSE SHAPING*

- The Nyquist Spectrum Signal Pulse has a Weaker Rate Tone when Compared to the Rectangular Pulse Shape of BPSK and is also More Sensitive to Receiver Parameters

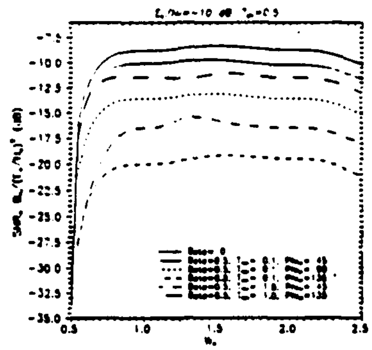


VIEWGRAPH #7

DELAY AND MULTIPLY: *MULTIPATH CHANNEL*



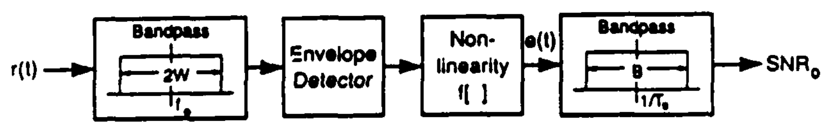
- For Fixed Channel Characteristics Severe Performance Degradations may Occur



- If the Phase of the Secondary Path, $\phi_m = T_m\,\omega_c$, is taken to be Uniform on [0, 2π], then the Delay and Multiply Receiver Performs as if no Multipath were Present and Received Signal Power is the Noncoherent Sum

VIEWGRAPH #8

POWER SERIES NONLINEARITY



- Nonlinearity Output

$$e(t) = \sum_k \alpha_k \sum_{i=1}^{k} \binom{k}{i} \sum_{j=1}^{i} A^{2j} \binom{2i}{2j} \hat{d}^{2j}(t)\hat{n}_c^{2i-2j}(t)\hat{n}_s^{2k-2i}(t)$$

where

$$d(t) = \sum_{k=-\infty}^{\infty} a_k p(t - kT_s) \quad \text{(signal modulation)}$$

- Output SNR Expression

$$SNR_o = \frac{|c_1|^2}{B\left[S_e\left(\frac{1}{T_s}\right)\right]}$$

- Background Noise for Low Input SNR (E_s/N_o)

$$R_e(\tau) = \sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_i\alpha_j \sum_{k=0}^{i}\sum_{n=0}^{j} \binom{i}{k}\binom{j}{n} R_n(\tau; 2k, 2n) \cdot R_n(\tau; 2i-2k, 2j-2n)$$

VIEWGRAPH #9

when you go to the more bandlimited signal. This model [VIEWGRAPH #10] now consists of a bandpass filter which is just the RF version of what we were talking about before, then an envelope detector and then some nonlinearity – in this case the nonlinearity is always going to be some power function or power series or just a single power function – and then the bandpass filter again. I guess I probably didn't mention before that this is the observation system out here, and in practice this would be replaced by some spectral estimation technique thrown out there. The bandpass filter is just a convenient model so you can say you're observing bandwidth B at some center frequency.

An expansion for $f[\ ]$ is the power series nonlinearity – we see here that the power series has coefficients α_k. If you expand this envelope out here in terms of signal and inphase and quadrature noise terms using the binomial expansion theorem you get something like this where the $d(t)$ is the modulated pulse stream or the pulse stream which, in this case, is going to have an amplitude weight a_k and pulse shape $p(t)$. The SNR expression, which corresponds to this, will again be the Fourier coefficient, the fundamental rate tone and then some noise spectrum which comes from this $e(t)$. If you extract out just the noise-related terms – this is the expression for the noise process right at this point ($e(t)$) in terms of the autocorrelation function of the noise and weightings with the α_k's from the power series.

This is the way you go about calculating the rate tone itself in terms of the Fourier series again. [VIEWGRAPH #11] This is the expression for $e_s(t)$, the part of the power series output which will contain the rate tone once you extract the Fourier coefficient. The important thing to note here is getting this Fourier coefficient from this expression: these are moments, higher order moments, of the noise process. So that the rate line actually is going to depend on not only the signal component coming in, but the noise is also going to affect the rate tone itself which actually complicates matters. It also makes some undesirable things happen but it does allow you to recover rate tones in bandlimited signals.

The way you get some of these things calculated is that first of all you have these higher order moments of the information signal itself to calculate. The way that we've looked at that is using cumulants of $d(t)$. This is the expression for the cumulants of the data signal in terms of the cumulants of the amplitude weights put on the data signal, and then powers of the pulse shape function itself. So this is where the idea of having a varied symbol probability distribution comes in. Before we go look at that though, one thing that we did was just look at very narrow bandwidth signals just to see how this works with narrowing up the bandwidth.

First of all we know when the bandwidth gets less than $\frac{1}{2\text{ symbol rate}}$ there's no rate tone generated by a squaring device or delay and multiply circuits in general. [VIEWGRAPH #12] So you need to go to higher powers. The first thing we looked at was 4th power. If you look at a power series and have all those terms collected together, or just even a 4th power circuit by itself, you have a problem in that you get nulls generated in the signal-to-noise ratio performance at various operating conditions. So you get some very unpredictable results. You might have one particular input signal-to-noise ratio, one pulse shape and it might work fine. Then you might have a parameter change and find you drop into a deep null, because you have cancellation effects. The fact is that the rate tone

RATE-LINE AMPLITUDE

- Rate-Line Fourier Coefficient

$$c_1 = \frac{1}{T}\int_0^T e_s(t)e^{j2\pi t/T_s}\,dt$$

where

$$e_s(t) = \sum_{k=1}^{N} \alpha_k \sum_{i=1}^{k} \binom{k}{i} \sum_{j=1}^{i} A^{2j}\binom{2i}{2j} E\{\hat{d}^{2j}(t)\} \cdot M_{2i-2j}M_{2k-2i}$$

- Rate-Line Depends on Signal and Noise
- The Moments $E\{\hat{d}^n(t)\}$ can be Found from the Cumulants of d(t) via a Recursive Relationship

  The Cumulants are given by

$$\lambda_d(n) = \lambda_a(n)A^n \sum_{k=-\infty}^{\infty} \hat{p}^n(t - kT_s) \quad (\text{Note: } \lambda_d(1) = m_d(t),\ \lambda_d(2) = \sigma_d^2(t))$$

- The Cumulants $\lambda_a(n)$ Depend on the Symbol Probability Distribution

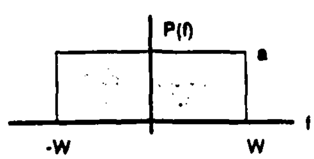
VIEWGRAPH #10

NARROW BANDWIDTH SIGNALS

- For Signals with Pulse Bandwidth Less Than $1/(2T_s)$ the Delay and Multiply Receiver Fails (also the Square-Law Receiver)
- In General a Power Series Nonlinearity may be used to Generate a Rate Tone, but SNR_o Nulls may Result
- For Binary Amplitude Weighting and Low Es/No a Power-Law Receiver may be used to Generate Rate Tones from Bandlimited Signals
- Let 2k be the Detector Power-Law and W the Signal Lowpass Pulse Bandwidth, then we can Choose k as Follows

$$2(k-1)W \le \frac{1}{T_s} \le 2kW$$

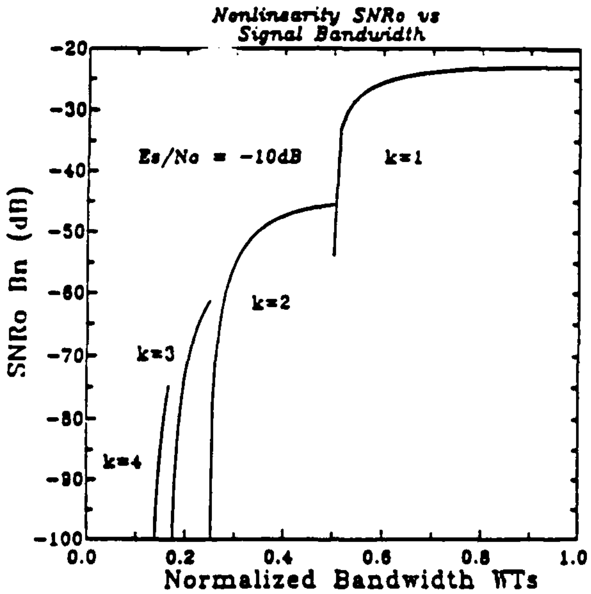
- For Analysis Purposes a Pulse with Rectangular Spectrum as Shown Below was used



VIEWGRAPH #11

NARROW BANDWIDTH SIGNALS CONT.

- Single Power-Law Nonlinearity SNR Performance



VIEWGRAPH #12

COMBINED PULSE SHAPING AND AMPLITUDE WEIGHTING

- To Begin with Suppose that the Symbol Amplitude Weights are Drawn from a Continuous Distribution
- For $\{a_k\}$ a Sequence of Zero Mean, Unit Variance Gaussian R.V.s

$$\lambda_a(n) = \begin{cases} 1, & n = 2 \\ 0, & \text{otherwise} \end{cases}$$

- The Moments are Found to be

$$E\{\hat{d}^n(t)\} = (n)!!A_n \left[\frac{1}{T_s} \sum_{k=-\infty}^{\infty} P_2\left(\frac{k}{T_s}\right) e^{j2\pi kt/T_s}\right]^{n/2}, \quad n \text{ even}$$

where

$$P_2(f) = \mathcal{F}\{\hat{p}^2(t)\}$$

- If p(t) is Bandlimited to $1/(2T_s)$ then $P_2(k/T_s) = 0$ for $k \neq 0$ then the Moments are Constant with Respect to Time and there is **No Rate Tone**

VIEWGRAPH #13

is now dependent upon not just the signal parameters but the noise as well. And of course that is controlled by shaping that goes on in the receiver. So you get unpredictable things that can happen, so that's something to be aware of.

Just to go to a simpler case of the power series, though, we've just worked with a low $\frac{E_s}{N_0}$ here, consider a single power law device, not a sum of power law devices. A power law design rule that was derived says that if you choose the power law to be $2k$, where k is an integer, then $\frac{1}{T_s}$ falls between these two limits, that's an approximately optimum power law to use. For analysis purposes, a bandlimited spectrum, very ideal type of model, was used to get something that you could analytically calculate.

So when you go to plot out the performance you see I guess what you'd expect to see. [VIEWGRAPH #13] You see that with $k = 1$ or a squaring circuit, the receiver fails when you get bandlimited to half the symbol rate. So then you'd switch over to a 4th power device, and it fails when you get to a quarter, and then 6th power and then 8th power. As you're continuing to use higher and higher powers, the noise is getting increased by those corresponding powers, and you're continuing to drop down lower and lower in your output signal-to-noise ratios. So whether or not you can actually pull something out is questionable, but theoretically the rate line does exist. It's there in an ideal sense. So this would be sort of a rule for truly bandlimited signals, how you would go and switch to different powers as the bandwidth decreases.

OK, the last thing I want to look at is just putting amplitude weighting on. [VIEWGRAPH #14] The previous signal was just binary amplitude weighting. So now what we're going to do is let the symbol weights take on some arbitrary distribution. Assume that the symbol weights take on a unit mean or zero mean unit variance Gaussian distribution. That's just for modeling purposes only. You go through the analysis of the moments and here's the cumulants for that amplitude distribution. You go through the analysis of the moments of the data sequence itself, and you find out that if you bandlimit your pulse shape to 1/2 the symbol rate, the moments end up being constants. So there's no rate tone generated, it's just like a DC component. You're not going to generate any rate tone. So that in general for a Gaussian amplitude weighting and bandlimited to 1/2 the symbol rate, it doesn't matter what power law device you use. You get nothing.

That's not a practical system so what we considered doing was quantizing, making the amplitude distributions be a quantized version of something that might work better than just the binary case. [VIEWGRAPH #15] So just to come up with the next higher order receiver from the delay and multiply or squaring circuit, we just assumed we had a 4th power nonlinearity, and further assumed that the receiver signal spectrum was rectangular with some lowpass bandwidth W, so we could just have another parameter to vary. The rate tone expression works out to look like this. It involves the pulse shape functions in the frequency domain. Note that $P_2(f)$ is the convolution between $p(f)$ with itself and then $P_4(f)$ would be a 4th order convolution, so this is broader spectrum than $P_2(f)$. The second term in c_1 also involves the 4th cumulant. So if you could ... well, let me go further before I say what I was going to say. This is just an example of some of the quantizations we could look at. This would just be putting a uniform weighting, this would be like doing some M-ary amplitude distributions. But if

PULSE AND AMPLITUDE WEIGHTING

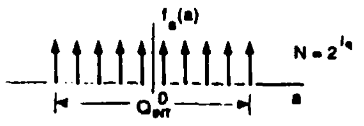
- Assume the Receiver consists of a Fourth Power Nonlinearity

$$f(\cdot) = [\cdot]^4$$

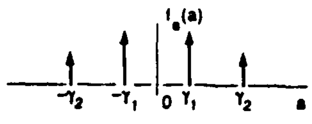
- Further Assume that the Received Signal Spectrum is Rectangular with Lowpass Bandwidth W
- The Rate Tone Expression is

$$c_1 = \left|8\sigma^2 A^2 + 6A^4\right| \frac{1}{T_s} P_2\left(\frac{1}{T_s}\right) + A^4 \lambda_a(4) \frac{1}{T_s} P_4\left(\frac{1}{T_s}\right)$$

- Quantized Symbol Amplitudes



Uniform Quantizer



Nonuniform Quantizer

- The Cumulants are

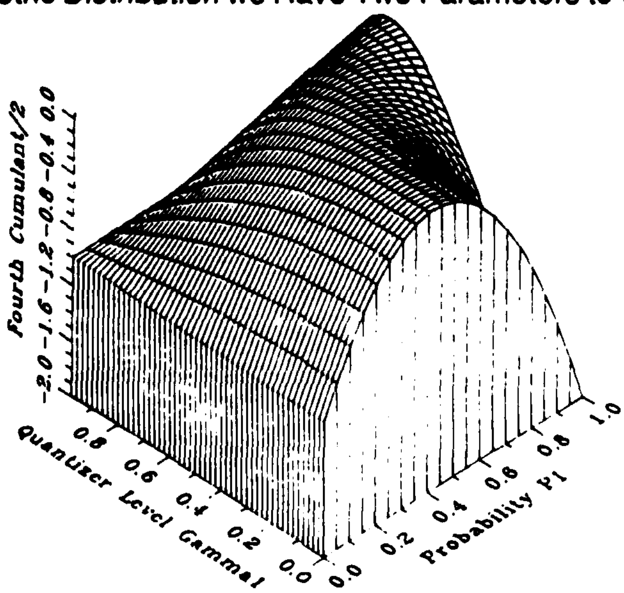
$$\lambda_a(2) = 2\sum_{i=1}^{2^{l_q-1}} p_i \gamma_i^2$$

$$\lambda_a(4) = 2\sum_{i=1}^{2^{l_q-1}} p_i \gamma_i^4 - 3[\lambda_a(2)]^2$$

VIEWGRAPH #14

OPTIMIZED FOUR LEVEL SIGNAL

- Choose the Distribution so $\lambda_a(2) = 1$ and $\lambda_a(4) = 0$
- Using a Symmetric Distribution we Have Two Parameters to Choose; γ_1 and p_1



VIEWGRAPH #15

you'd let it be non-uniform amplitude distribution, perhaps just a quantized version of a Gaussian amplitude distribution, maybe that would work in a limited sense like the Gaussian distribution, but yet have hopes of being realizable. The cumulants for a general non-uniform amplitude distribution like this work out to these two equations. Here's the second and the fourth cumulant. The fourth cumulant is the one we're interested in plugging into our c_1 expression. This one we could take care of by letting this pulse shape be bandlimited and then this term would drop out of c_1 for the rate tone. So this is the one (eliminate $\lambda_a(4)$) we expanded this one out for just a 4th four-level quantizer. You end up having then two parameters to vary. You have the probability weights – if you have a four-level quantizer you have the weights of each of the four levels, but we made it symmetrical so there were two probabilities and two amplitude weights. Well, the probabilities have to sum up to 1. [VIEWGRAPH #16] So if you vary the probability of one of those amplitude weights and then let the quantizer level vary, you get this surface for the fourth cumulant. There's a region in here which is sort of horseshoe shaped which corresponds to the fourth cumulant being zero, so that would be the place you'd want to set amplitude weights and quantizer probabilities so that you could get that fourth cumulant to go to zero. Then what would happen is that a 4th power device with a bandlimited signal would fail to generate a rate tone. It's a bunch of conditions and maybe it doesn't seem very realistic, but it's just looking at ways of approaching something where you could wipe everything out.

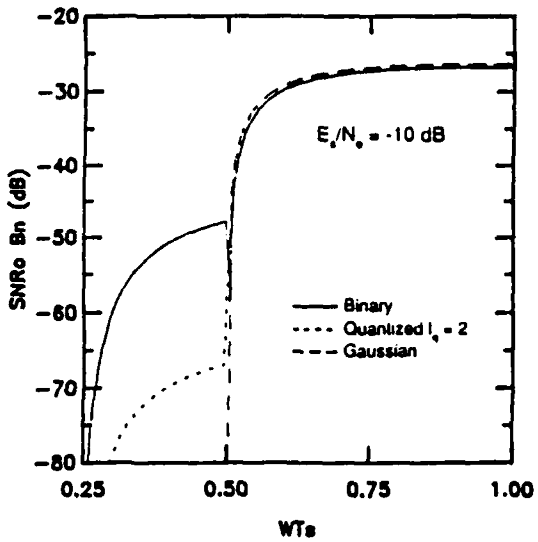
Maybe something that's more practical is just not trying to be fancy with your quantizer, just making a quantized Gaussian version for your amplitude distribution and just seeing once you vary the bandwidth what happens. [VIEWGRAPH #17] As you get down to the half bandwidth point, the ideal Gaussian goes away. The binary has a null and then comes back up again because this is a 4th power device. Remember I said it has cancellation points. The quantized signal with just 2^2 or four levels here, you get a considerable improvement as far as suppressing the rate tone generated. This plot over here then shows extending to more levels and using a wider quantization interval. Some of the options you can get as far as suppressing the rate tone amplitude are shown here. Here you're just using four levels so that when you spread them out further over a bigger amplitude region you actually come back up again, but if you used more levels you continue to drop down, which is kind of like approaching the Gaussian again, then, by letting the spread (quantization interval) get big and the number of levels get large.

These are some of the conclusions. [VIEWGRAPH #18] A lot of the points made were on the negative aspects, I guess, how not to make it work. Rate tone detectability can be reduced by bandlimiting. We looked at, in particular, the case of the sinusoidal roll-off; and what that does is that the insensitivity of the parameters were increased, which means it would make it harder in general to search to find something. I didn't show you this slide [VIEWGRAPH #9], but multipath channels with fixed channel parameters you can get conditions where you wipe everything out. However if you have a channel where the phase is uniformly distributed, it turns out that it looks like the multipath just is a non-coherent power summing right at the receiver. So you're not really in bad shape there.

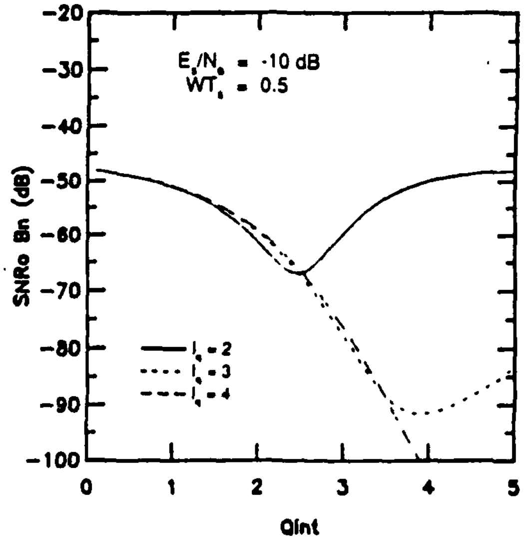
For bandlimited signals you want to use the lowest power law which gives a rate tone,

FOURTH POWER CIRCUIT PERFORMANCE

- SNRo versus Signal Bandwidth



- SNRo versus Quantizer Range for Uniform Quantization of a Gaussian Distribution



VIEWGRAPH #16

CONCLUSIONS

- Rate Tone Detectability by a Delay and Multiply Circuit can be Reduced by Using a Bandlimiting Pulse Shape Function
- A Pulse Spectrum with Sinusoidal Roll-off Increases the Sensitivity of the Delay and Multiply Receiver Prefilter Bandwidth
- Multipath Channels with Fixed Channel Parameters can Seriously Degrade the Performance of a Delay and Multiply Receiver
- For Bandlimited Signals using the Lowest Power Law which gives a Rate Tone is Approximately Optimum for Low Input SNRs
- Combined Higher Order Nonlinearities Suffer From Rate-Tone Cancellation for Certain Combinations of Receiver and Signal Parameters
- Amplitude Distributions can be Optimized for Bandlimited signals to Defeat Rate Tone Generation Circuits
- A Simple Four-Level Amplitude Signal can be Used to Defeat a Fourth-Power Receiver

VIEWGRAPH #17

UNRESOLVED ISSUES

- ❑ Practical Hardware Realization
- ❑ Communicator Performance Degradation
- ❑ Spectrum Resolution Limits
- ❑ Others

VIEWGRAPH #18

STATISTICAL PATTERN RECOGNITION VERSUS MODEL-BASED APPROACHES TO SIGNAL CLASSIFICATION

SLIDE 1

TCI TECHNOLOGY FOR COMMUNICATIONS INTERNATIONAL

MM-101-89

STATISTICAL PATTERN RECOGNITION HAS BEEN USED FOR...

- MODULATION IDENTIFICATION
- MORSE DECODING
- LANGUAGE IDENTIFICATION
- ATMOSPHERIC ANALYSIS
- BATTLEFIELD TARGET IDENTIFICATION
- INTRUSION DETECTION
- SPEAKER RECOGNITION
- HANDWRITING AUTHENTICATION

SLIDE 2

TCI TECHNOLOGY FOR COMMUNICATIONS INTERNATIONAL

MM-101-89

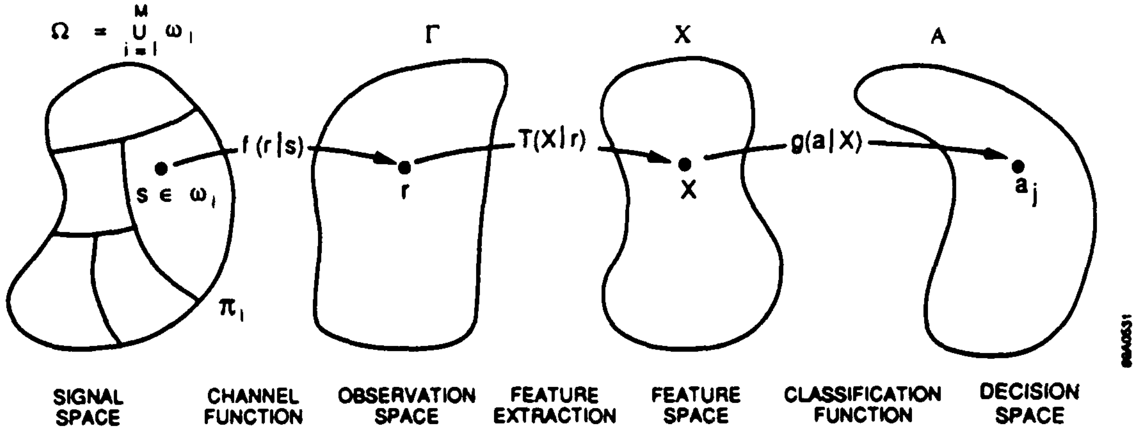
ments of Signal Detection. I've generalized the picture somewhat to show some additional features. The conceptual idea here is that we have a signal space which I've labelled Ω, which is divided into subclasses labelled with ω (lower case omega), and that we have some procedure nature, or an adversary selects a subclass and selects a signal within a subclass, and there is some probabilistic mapping from that signal into an observation space to produce a received signal. An element of a space called γ, I've labelled it r; in general it's an analog waveform, usually at the IF of a receiver. We then proceed to have two mappings which result in a decision. The first mapping consists of a feature extraction and the second mapping consists of a classifier decision classification function. The result of all three of those probabilistic mappings – the channel function, the feature extraction and *the classification function – once they're* all specified, we can determine conditional probabilities of various class decisions conditioned on various signal classes occurring, and these define the elements of the confusion matrix. The confusion matrix is a square matrix of conditional decision probabilities. You'll see that the feature extraction classification function enters in the integrand in the form of a product and the complexity of the decision can be put either entirely in the feature extraction operation or entirely in the classification function, or it can be divided between the two. It doesn't really have any effect on the overall performance. The probability of error is a means of summarizing what is in the confusion matrix and it consists of a weighting of the diagonal elements with the prior probabilities in summing them, subtracting from 1.

[SLIDE 4] Showing the same figure we have an input signal which is either an IF or a baseband audio or video signal which is observed for a finite amount of time. We extract some number of features and produce a feature vector x which is finite dimensional. Then that gets passed to a pattern classifier which computes a set of discriminate functions, one for each class. If there are m classes, there are m discriminate functions. If the discriminate functions are interpreted as distances, then we'll choose the smallest discriminate function and report its index. The index will give us – depending on what resolution we're using and the definition of what constitutes a class – a signal class, modulation, signal type. Within the class it'll give us an I.D., and the extreme case will do radio fingerprinting and will identify a particular transmitter down to the serial number of the transmitter.

In a system where we're doing all of these operations at the various levels, we'll partition it so that we'll do the coarse resolution decision-making first and then the fine decision-making later, and the system will be hierarchical. This picture shows the features inside the feature extraction being computed in parallel; in practice they're generally not that nicely divided. Some features will be computed from other features, so we have a dependency. For example we may compute a spectrum, then from the spectrum we'll compute certain features, and then from those features we'll compute other derivative features so that we don't have independent parallel computation of each feature.

I'm not addressing sequential decision-making and I'm not addressing the difficulty of designing decision trees to implement sequential decision-making. I'm assuming here that we're talking about a classifier that computes all the features that are going to be computed and then looks at the entire vector of features and makes a decision.

SIGNAL CLASSIFICATION PROCESS USING STATISTICAL PATTERN RECOGNITION



PERFORMANCE MEASURES:

CONFUSION MATRIX: $C = [c_{ij}],\quad c_{ij} = P(a_j|\omega_i) = \int_{\Gamma}\int_{X} f(r|\omega_i)\, T(X|r)\, g(a_j|X)\, dX\, dr$

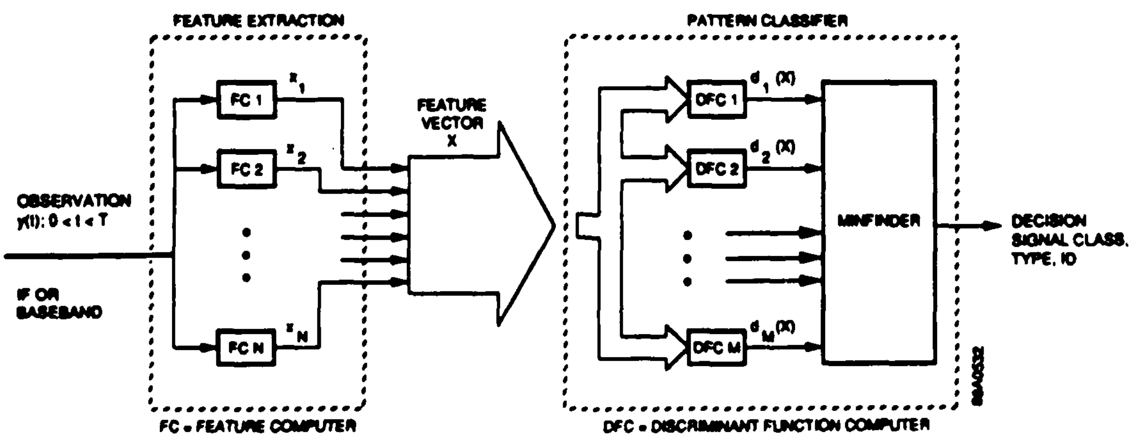
PROBABILITY OF ERROR: $P_e = 1 - P_c,\quad P_c = P(a_j = \omega_i) = \sum_{i=1}^{M} \pi_i c_{ii}$

SLIDE 3

TCI TECHNOLOGY FOR COMMUNICATIONS INTERNATIONAL

MM-101-89

PATTERN RECOGNITION APPROACH TO SIGNAL CLASSIFICATION

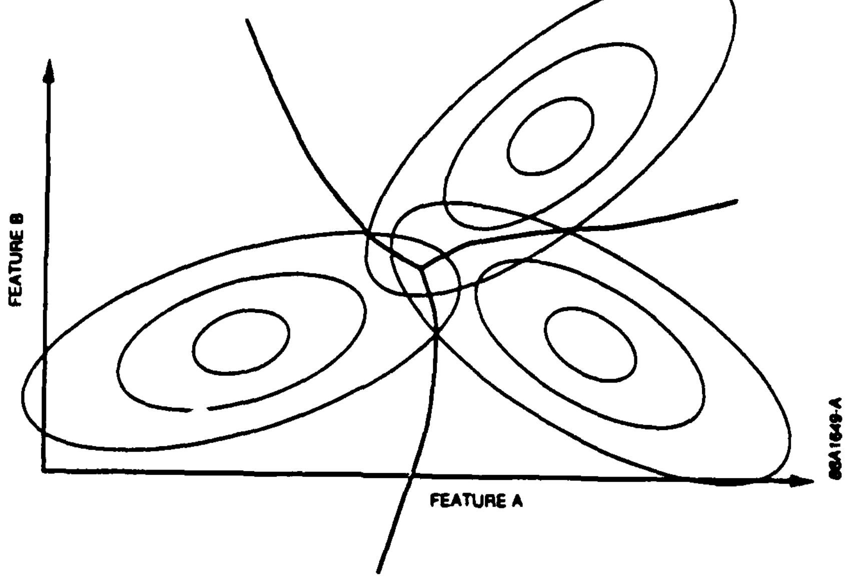


SLIDE 4

TCI TECHNOLOGY FOR COMMUNICATIONS INTERNATIONAL

MM-101-89

PROBABILITY CONTOURS

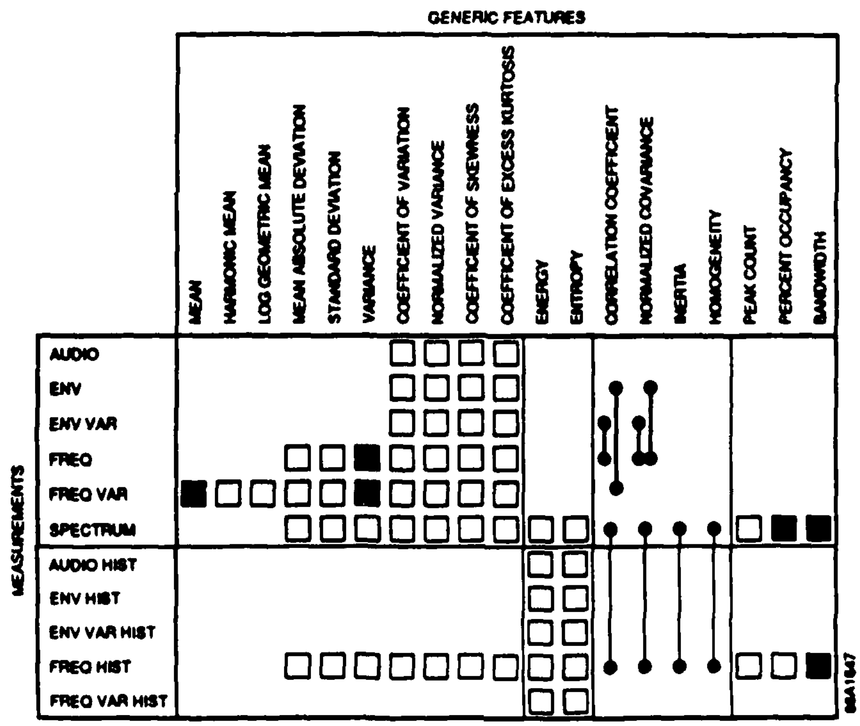


TECHNOLOGY FOR COMMUNICATIONS INTERNATIONAL

SLIDE 5

MM-101-89

SOME TYPICAL FEATURES



TECHNOLOGY FOR COMMUNICATIONS INTERNATIONAL

SLIDE 6

MM-101-89

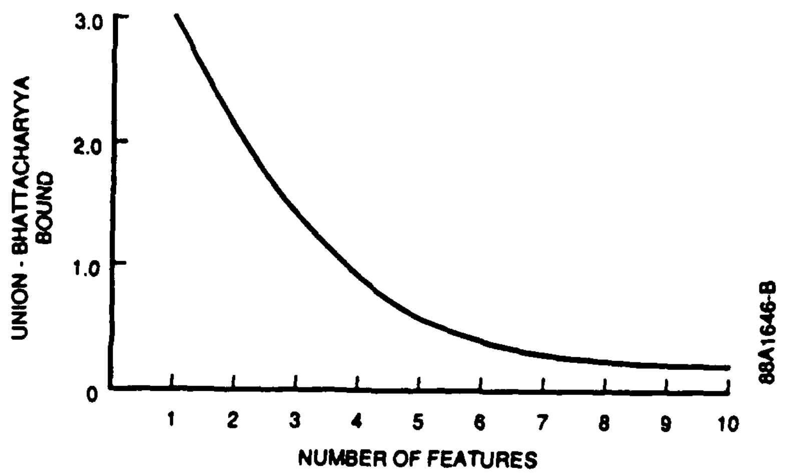
The job of designing a classifier consists of determining the boundaries between decision regions that will separate the classes. Boundaries can be straight lines in the case of a linear classifier. It can be segments of lines connected in the case of a piecewise linear classifier such as a nearest-neighbor classifier. The boundaries can be constrained to be hyperplanes parallell to the all axes except one, and orthogonal to that one axis. That would be the case of a sequential classifier which is making its decision on the basis of thresholding individual feature values. The boundaries can be curved surfaces in the case of quadratic and higher order polynomial classifiers. So there's a variety of options available in determining the type of classifier and also the choice of features that are used.

[SLIDE 7] In determining features the method that's used is entirely empirical. One simply constructs a large library of candidate features, computes them from the collected samples of the signature data, and then uses an automatic feature selection program to determine optimum subsets. We might start with 100 features, typically, and then determine the optimum singleton feature, that is, the best single feature. We'll determine the performance of a classifier using that feature, then we'll repeat that to determine the best pair of features, the best triplet of features, the best quartuple of features. At each stage we will find the performance of that feature set of that particular size, and we'll watch as the performance declines as the number of features increases, determine a suitable operating point on this curve, and that will determine the size of the feature set.

Now the particular features that are chosen, as you can see from this figure, are all ad hoc statistics, descriptive statistics. We attempt in constructing this list of candidate features to choose features that have certain properties that we think are going to be useful *a priori*. For example it may happen that all the collected signature data was collected at signal-to-noise ratios between 10 and 20 dB. If a classifier were optimized on such data you'd expect it to perform well on signals between 10 and 20 dB signal-to-noise ratio. If you gave it a signal with a signal-to-noise ratio lower than 10 or higher than 20, you might expect it to fail and that's exactly what happens. Performance is not monotonic with signal-to-noise ratio unless we do something to cause it to be monotonic. What we might do to cause it to be monotonic with signal-to-noise ratio would be to require certain invariance properties, such as we might require the features to be invariant with respect to scale changes in amplitude of the input signal. We might require the features to be invariant with respect to frequency shifts. We can't guarantee that a receiver will be tuned precisely to the center frequency of the signals, so if the receiver is mistuned slightly we want the features to be invariant through that mistuning. [SLIDES 8,9]

Performance that is typically achieved might look like this. We have a confusion matrix here, where the numbers on the diagonal represent the probabilities of the various types of correct decisions; the numbers on the off diagonal represent the different kinds of error. We have more than a Type 1 and a Type 2 error, obviously. This is a multi-category decision problem. Each row of this matrix sums to unit probability. The objective, of course, is to design a confusion matrix that looks like the identity matrix, which has 1's down the diagonal. In practice you can't do this, but you can typically get numbers that are on the orders of 95%, 99% in some cases. Error probabilities would be in

P_E BOUND VERSUS SIZE OF FEATURE SET



SLIDE 7

EXAMPLE OF TYPICAL CLASSIFICATION PERFORMANCE ACHIEVED BY STATISTICAL PATTERN RECOGNITION METHOD

DECISION CLASS

TRUE CLASS	FM	AM	OOK	FSK	CW	PSK
FM:	.89	.08	.01	.02	.00	.00
AM:	.00	.97	.01	.01	.00	.01
OOK:	.03	.03	.93	.01	.00	.00
FSK:	.06	.03	.01	.90	.00	.00
CW:	.00	.00	.00	.00	1.00	.00
PSK:	.00	.00	.00	.03	.00	.97

SLIDE 8

TCI TECHNOLOGY FOR COMMUNICATIONS INTERNATIONAL
MM-101-89

the range of 5%; 10% representing a fairly easy design number, 1% representing a difficult design number. It's easy to fool oneself if you work with a small database because you can always get 100% on any small database. If you test your classifier design on the same data that you designed it on, it's easy to overfit the data and get perfect performance which won't extrapolate to the real world.

[SLIDE 9] If we show error probability versus signal-to-noise ratio we get curves that look like this. However, the curve may not continue to decline as signal-to-noise ratio is increased because certain features are nonlinear functions of the data, and the distribution of that feature will translate through the feature space. And in translating it will eventually shift outside the, or cross over a, decision boundary. Once the probability mass has crossed a decision boundary then that signal is misclassified, even though it may have very high signal-to-noise ratio.

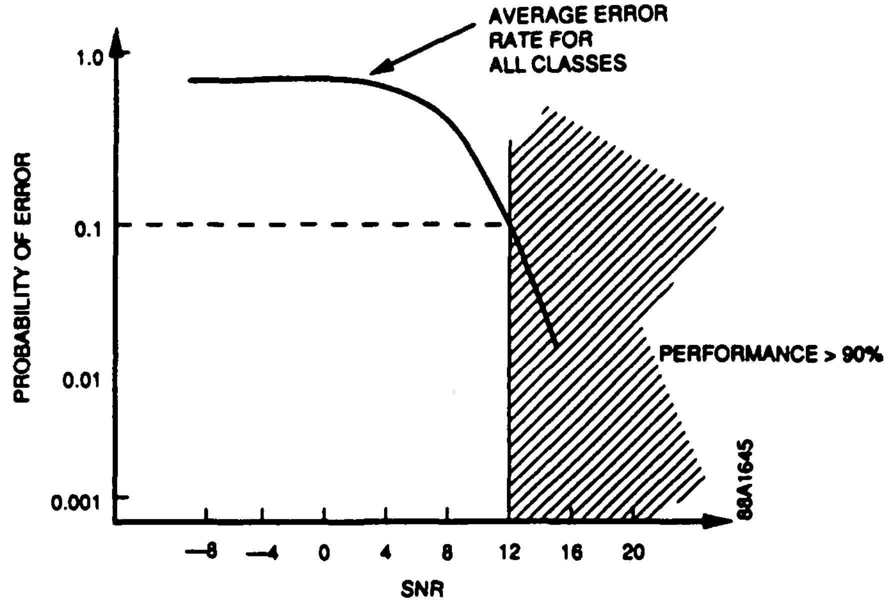
So those are some of the practical problems connected with statistical pattern recognition. I haven't mentioned the size of the databases that are typically used. We might collect sufficient data to compute maybe 3,000-10,000 feature vectors per class, and a typical problem could have anywhere from 10 to 20 or 30 classes. In some cases the numbers are much higher. We have the luxury of unlimited computer time to design these classifiers. So we can let the computer do its number-crunching for several days at a time to come up with optimum selections, subsets of features, optimum classifier weights and so on.

[SLIDE 10] Now I'd like to shift to an altogether different approach to signal recognition. This is based on a model of how the signals are generated and tries to exploit information that's available in the signals. Now we don't always have good models, but in the case of communication and radar signals that are being transmitted the modulations are well specified by the system design engineer. The waveforms have specific structure that's put in them to make them easy for a communication receiver to demodulate, and there's no reason why an intercept receiver couldn't use that same structure to its advantage.

Here we have an input from an antenna, presumably a receiver front end, which I've called "signal preprocessor" which consists of various mixing operations and bandpass filtering, producing a signal at either IF or baseband audio or video signal. The signal processing operations we would like to do would be to compute some type of generalized likelihood function for every possible signal class. We'd like that likelihood function to be based on some knowledge of what the noise characteristics are in the particular frequency band of interest, and also some knowledge of what the channel is doing to the signal. The channel model would include the receiver front end in this case as well.

We can set this up as a multiple category hypothesis problem. Hypothesis H_0 would represent signal signal absent; we're looking at noise only. And we'll have some number M of hypotheses that represent the different signal types that might be present. As discriminant functions we will take the *a posteriori* probability of the hypothesis, given the observed waveform $y(t)$. I'm going to normalize all of these by the conditional probability of the Null Hypothesis to get ratios of these *a posteriori* probabilities. My decision rule would be to choose the J^{th} hypothesis if the J^{th} discriminant function dominates all the other discriminant functions. The discrimi-

P_E VERSUS SIGNAL-TO-NOISE RATIO

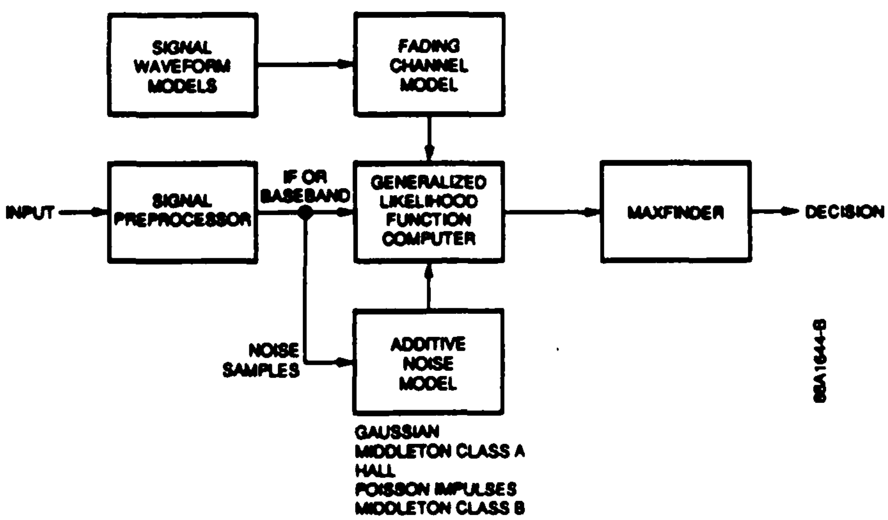


SLIDE 9

TCI TECHNOLOGY FOR COMMUNICATIONS INTERNATIONAL

MM-101-89

MODEL-BASED SIGNAL DETECTION AND CLASSIFICATION



SLIDE 10

TCI TECHNOLOGY FOR COMMUNICATIONS INTERNATIONAL

MM-101-89

JOINT DETECTION AND CLASSIFICATION

HYPOTHESES: H_0 : SIGNAL ABSENT, AWGN ONLY

H_i : SIGNAL TYPE i PRESENT, $i = 1,2,...,M$

DISCRIMINANT FUNCTIONS:

$$d_0(y(t)) = 1$$

$$d_1(y(t)) = \frac{P(H_1 \mid y(t))}{P(H_0 \mid y(t))}$$

$$\vdots$$

$$d_M(y(t)) = \frac{P(H_M \mid y(t))}{P(H_0 \mid y(t))}$$

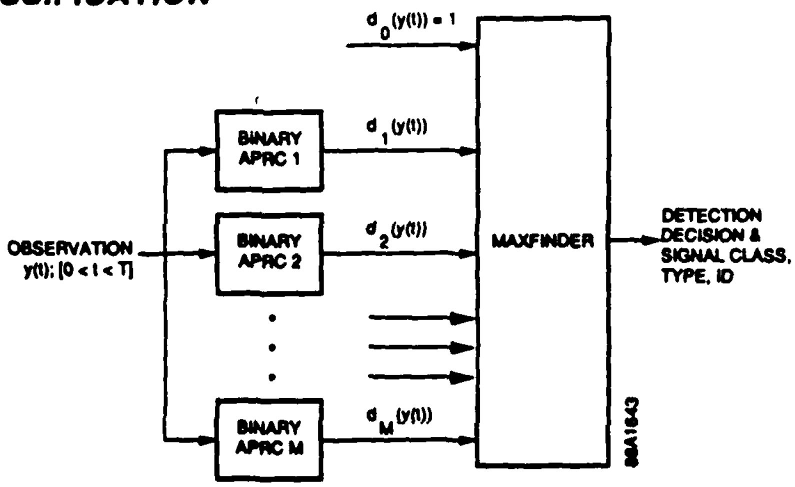
RULE: "CHOOSE H_j IF $d_j(y(t)) \geq d_k(y(t))$ FOR $k \neq j$"

TCI TECHNOLOGY FOR COMMUNICATIONS INTERNATIONAL

SLIDE 11

MM-101-89

SIMULTANEOUS DETECTION AND CLASSIFICATION



APRC = A POSTERIORI PROBABILITY RATIO COMPUTER

$$d_k(y(t)) = \frac{P(\omega_k \mid y(t))}{P(\omega_0 \mid y(t))} = L_k(y(t)) \frac{\pi_k}{\pi_0}$$

TCI TECHNOLOGY FOR COMMUNICATIONS INTERNATIONAL

SLIDE 12

MM-101-89

nant function for the Null Hypothesis is unity by virtue of the normalization.

[SLIDE 12] A block diagram, which says the same thing as the equations, looks like this. We have these little boxes that I call binary *a posteriori* probability ratio computers which simply compute these probability ratios, followed by a max taker that finds the largest discriminant function and reports the index. You'll notice that these discriminant functions are related to likelihood functions, testing each signal class against the Null Hypothesis. The likelihood ratios have to be multiplied by the ratio of the priors, $\frac{\pi_k}{\pi_0}$, in order to establish the connection.

Now you might think of this as a pattern classifier where we have exactly $M + 1$ features, the features are these discriminant functions. The feature computation is carried out in these [UNINTERPRETABLE, STATIC INTERFERENCE] probability ratio computers. The classifier is simply the max finder. So all the complexity of feature [UNINTERPRETABLE, STATIC INTERFERENCE] computation is in the calculation of these likelihood functionals. Now we need to look at what the complexity is here because in a real problem an intercept receiver, unlike a communication receiver, is faced with a far greater number of unknown parameters. [LONG PAUSE, POSSIBLE DROPOUT]

[SLIDE 13] Now in statistical decision theory we have hypothesis tests that are appropriate to each of these three cases. In the first case where we have known parameters, we have simple hypothesis tests. In the second case where we have unknown parameters with known probability distribution, we have composite hypothesis tests. And in the third case where we have unknown parameters with unknown probability distributions, we have generalized hypothesis tests (sometimes called maximum likelihood tests). What we can do is to combine the features of all three into a single test. What I've shown here is a signal that's parameterized with two parameter vectors, which I've labelled $\underline{\theta}_1$ and $\underline{\theta}_2$. The first parameter vector, $\underline{\theta}_1$, contains all the parameters that have a known probability distribution; and the second vector, $\underline{\theta}_2$, has all the parameters that have an unknown probability distribution. With respect to the first set of parameters I'm going to integrate them out as nuisance parameters in this integral, $d\underline{\theta}_1$. With respect to the second class of parameters I'm going to take a maximization, just as I would in a maximum likelihood estimation or a generalized likelihood ratio test.

The function $p(\underline{\theta}_1)$ is the joint density function of all the parameters in this vector, so this is a multivariate.

[SLIDE 14] A processor that implements these discriminant functions looks like this. This is a literal interpretation of the equation, and I'm not proposing that someone build this thing as I've drawn it. This is a brute force, this is a conceptual block diagram. We can think about it and talk about it and maybe even compute its performance. There obviously are a number of simplifications one would have to make.

What we have here is a processor which, with respect to $\underline{\theta}_1$, in the upper block θ_1 varies in each of these processing chains. We have a matched filter, an energy subtraction which normalizes energy, exponentiation, and then we multiply in a prior. Now I'm going to replace the integral with a sum, and I'm going to quantize the parameter space over $\underline{\theta}_1$ in such a way that I can make this replacement of an integral by a sum. I'm going to hold the vector $\underline{\theta}_2$ constant within this box and only vary θ_1. Each input into the summation

HYBRID GENERALIZED COMPOSITE TEST FOR SIGNALS WITH UNKNOWN PARAMETERS

LET: $y(t) = s_k(t, \underline{\theta}_1^{(k)}, \underline{\theta}_2^{(k)}) + n(t)$

WHERE:

$s_k(t, \underline{\theta}_1^{(k)}, \underline{\theta}_2^{(k)})$	IS A CLASS k SIGNAL WITH PARAMETERS $\underline{\theta}_1^{(k)}, \underline{\theta}_2^{(k)}$
$p(\underline{\theta}_1^{(k)})$	IS KNOWN
$p(\underline{\theta}_2^{(k)})$	IS UNKNOWN
$n(t)$	IS AWGN

THEN:

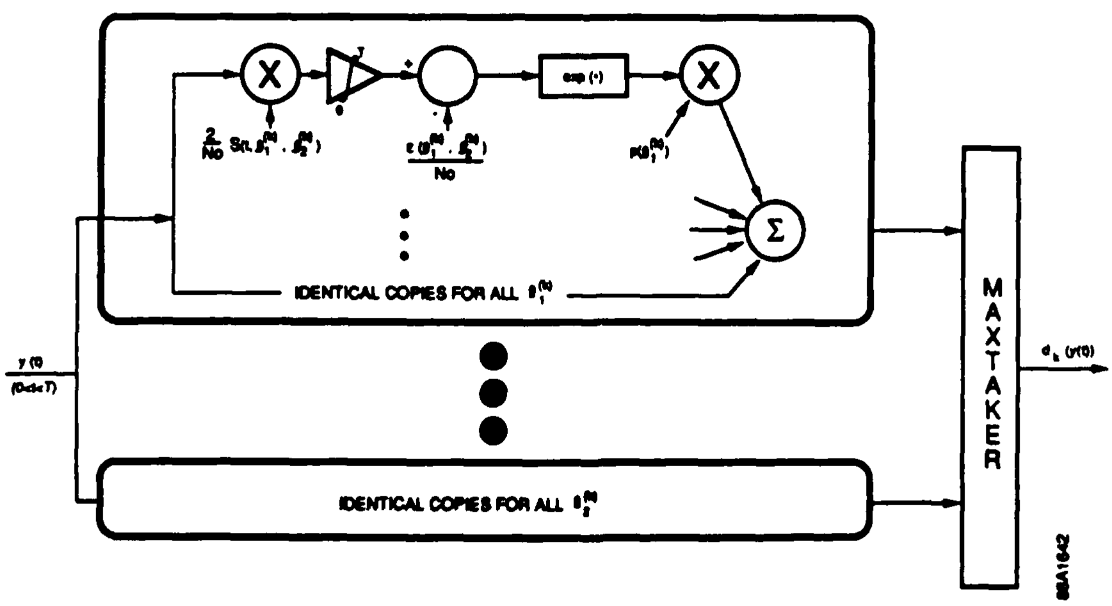
$$d_k(y(t)) = \max_{\underline{\theta}_2} \int e^{\left[\frac{2}{N_o}\int_o^T y(t)\, s_k(t, \underline{\theta}_1^{(k)}, \underline{\theta}_2^{(k)})\, dt - \frac{\mathcal{E}(\underline{\theta}_1^{(k)}, \underline{\theta}_2^{(k)})}{N_o}\right]} p(\underline{\theta}_1^{(k)})\, d\underline{\theta}_1^{(k)}$$

SLIDE 13

TCI TECHNOLOGY FOR COMMUNICATIONS INTERNATIONAL

MM-101-89

PROCESSOR TO COMPUTE PROBABILITY RATIOS



SLIDE 14

TCI TECHNOLOGY FOR COMMUNICATIONS INTERNATIONAL

MM-101-89

device will represent a different value of $\underline{\theta}_1$. Then I'm going to replicate this box many times for different values of $\underline{\theta}_2$. So each replication of this box will have some different $\underline{\theta}_2$, and I'm going to quantize the $\underline{\theta}_2$ space. Finally I'm going to take a max over the outputs of all these boxes, and then that will give me the discriminant function.

What I have here is a high degree of parallelism. First of all the first box is replicated multiple times for different values of $\underline{\theta}_1$. Then the first box is replicated multiple times for different values of $\underline{\theta}_2$. Then finally this entire picture is replicated multiple times, once for every box that shows up labelled as binary *a posteriori* probability ratio computer. And we can go on because we could say that this processor is being applied to an observation that comes out of a channelized receiver, so this entire box could be replicated multiple times on each frequency and on each angle of arrival, if one wants to precede it with some kind of a beamformer with fixed beams. So we could easily have hierarchical, five layers of hierarchical parallelism here.

[SLIDE 15] At this point I want to summarize the differences between the model-based approach, that I've just described, and the statistical pattern recognition approach to signal classification. Statistical pattern recognition is statistical because it's working on samples; that is, the design of the decision function comes not from a theoretical model of what the waveforms look like, but rather it comes from collected signals that are mapped and then various empirical methods are applied. The classifier has to be optimized both with respect to the choice of features and with respect to the weights, the choice of classification function. Whatever informations are stored in those collected samples, that's the information that is used for the decision. If you do not collect a representative database then your design is not going to extrapolate well to the real world. The features that result are unrelated to the physical model. You can come up with very strange moments being used as features, ratios of moments that are used as features that have no physical meaning. Error rates that are typically achieved are in the 5% to 10% range, 1% at the best. The classifier is very easy to implement, the design procedures are automatic. It's a cookbook procedure basically. The feature calculations are very complex, the classification is generally very simple in terms of computation. There's no architectural parallelism to exploit. As a rule, all the features are quite different from each other and one has to write subroutines that compute each feature.

The model-based approach, on the other hand, has a different set of advantages and disadvantages. First of all you do need to have models of the signals and of the noise as they appear at the IF of the receiver. This knowledge or need for knowledge of a model offsets the need to go out into the field and collect data. On these problems frequently one is interested in signals that don't appear in nature. There may be certain military signals that simply never are seen or at least you hope they never are seen, but you must design a recognizer that will recognize those signals very fast if they ever appear. You can't collect signature data on those signals because they simply don't manifest themselves in the real world. So the only way you could possibly design a recognizer would be on the basis of some description of the waveform.

The model-based approach does not require training. It uses all the available information about the signals of interest. The likelihood features are related to the physi-

COMPARISON OF APPROACHES TO SIGNAL RECOGNITION

STATISTICAL PATTERN RECOGNITION	MODEL-BASED APPROACH
REQUIRES STATISTICALLY-VALID CALIBRATED DATA BASE	REQUIRES MODELS OF SIGNALS AND OF NOISE
REQUIRES WEIGHT-TRAINING EXERCISE	NO DATA OR TRAINING EXERCISE REQUIRED
USES ONLY THE INFORMATION STORED IN THE FEATURES EXTRACTED FROM THE DATA	USES ALL AVAILABLE INFORMATION ABOUT SIGNAL-OF-INTEREST
FEATURES UNRELATED TO PHYSICAL MODEL	LIKELIHOOD "FEATURES" HAVE MEANING RELATED TO PHYSICAL MODEL
5% TO 10% ERROR RATE TYPICAL	ERROR RATE UNDER 1%
RELATIVELY STRAIGHTFORWARD TO IMPLEMENT	COMPUTATIONALLY-INTENSIVE, BUT
SLOW DUE TO 1). LONG OBSERVATION TIME REQUIRED 2). COMPLEX FEATURE CALCULATIONS	SUITABLE FOR IMPLEMENTATION WITH HIGHLY PARALLEL ARCHITECTURES
LITTLE OR NO ARCHITECTURAL PARALLELISM TO EXPLOIT	INTEGRATED DETECTION AND RECOGNITION FUNCTIONS

TCI TECHNOLOGY FOR COMMUNICATIONS INTERNATIONAL

SLIDE 15

MM-101-89

OPTICS VS SUPERCOMPUTERS

COMPUTATIONAL POWER PER UNIT VOLUME		
PROCESSOR	FLOPS/CUIN	RELATIVE RATING (LOGARITHMIC)
OPTICAL: COUNTERPROPAGATING PROCESSOR	2×10^8	7.8
ESL TRIPLE PRODUCT PROCESSOR	8×10^7	7.3
DIGITAL: AT&T DSP 32C	3×10^7	6.8
THINKING MACHINES CM-2	1.3×10^4	3.6
NUMERIX 332	7×10^3	3.3
µVAX	27 x 10	1.3
VAX 11/780	4 x 10	0

SOURCE: A. HELLAND, ESL INCORPORATED

TCI TECHNOLOGY FOR COMMUNICATIONS INTERNATIONAL

SLIDE 16

MM-101-89

cal model. The error rate can be made as small as one wishes, subject to signal-to-noise ratio limitations. It's computationally intensive but there's a high degree of parallelism that could be exploited. The detection and recognition functions can be combined rather than having an energy detector which is not signal-specific, and then following that with a signal recognizer one could combine the detection and recognition in one step.

[SLIDE 16] I have three more slides. I'd like to talk about implementation of this method briefly. If we look at digital techniques, the digital techniques simply are not fast enough to handle the large number of parameters and the quantization of the parameter space that's required. If we look at optical techniques we discover that in terms of computational density, we can have more flops per cubic inch. The numbers that are available right now are on the order of 10^8 flops/in^3. The numbers that are in the foreseeable future are on the order of 10^{11} flops/in^3. This means that the analog optical processors have the potential for far greater computational density than the fastest supercomputers.

Now the frequent criticism of optical processing is that the dynamic range is small and the lack of input devices. The reply to those two criticisms are that high dynamic range is not required for carrying out the computations in one of these signal classifiers. Typically 4 bits of computation, 5 bits, 5-bit wordlength is sufficient. The dynamic ranges that are achievable are in the order of 30 dB or better, although generally not better than 40 dB. That's adequate. That's not adequate for doing filtering and reproducing high fidelity replicas of signals, but it is adequate for doing the analog computations necessary to implement these signal classification operations.

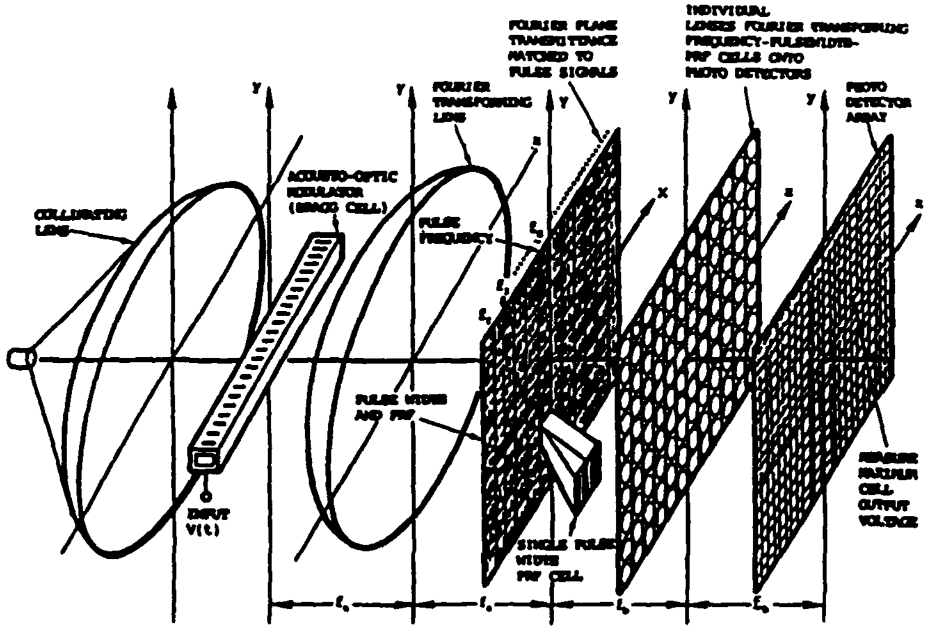
In response to the other criticism, the input device Bragg cells may not be the way to go but there are a variety of new spatial light modulators that are available that work on different physical principles, including liquid crystal devices. These may provide the storage capacity necessary to make a device work.

[SLIDE 17] I show here finally an optical bank of matched filters. This is one of the key ingredients in the model-based recognizer. This would be implemented with an array of lenslets, a transparency that encodes the matched filters for the signals for various parameter values, and an array of photodiodes. The array of lenslets: the current densities that are achievable are about 100 per centimeter of linear dimension or 10,000 lenslets/cm^2. We have no problem getting the photodiodes at densities greater than that. We could have several matched filters in a small region imaged by a single lenslet onto a single diode. The energy subtraction could be carried out after the photodiode.

[SLIDE 18] Here's a picture of a processor which is approximately 3 inches on a side. It's the Dove Prism Correlator that was built for the U.S. Army by Perkin-Elmer. It shows essentially the same picture as in this bank of matched filters. The light path in this device is folded up so that the device will fit nicely into a compact physical geometry.

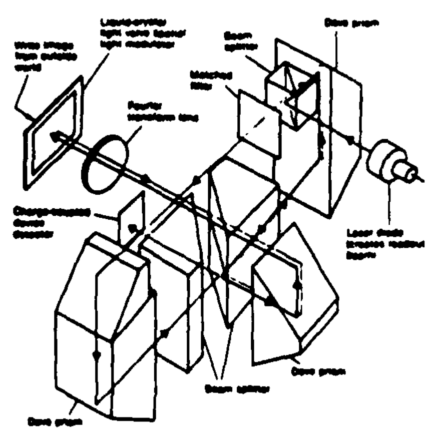
And we have here a laser diode which provides the input signal. The light path bounces out back here, spatially modulated by a reflection mode modulator, comes back into the device, comes back up through the transparency, and then winds its way back down and comes out onto the photodetector array. This device also was designed for missile guidance and designed to recognize cer-

OPTICAL MATCHED FILTER BANK



SOURCE: D.P. SULLIVAN, PhD THESIS, USC, 1987

U.S ARMY DOVE PRISM CORRELATOR OPTICAL PROCESSOR



SOURCE: T.E. BELL, *IEEE SPECTRUM*, AUGUST 1986

tain targets of interest, independent of the aspect angle and orientations. So it was doing image processing.

I'm proposing a replacement of this matched filter with an array of matched filters. This I think is where things will be headed in the next fifteen years. This area of investigation started fifteen years ago and it will probably be another fifteen years before we see this type of device with the high degree of parallelism that I'm anticipating.

WEBER: OK, we're running very late so we'll take a break now. [PAUSE, BREAK]

Here is Ed Satorius from JPL [PAUSE] and he's here!

ED SATORIUS: *Application of Neural Networks to Signal Sorting*

[SLIDE #0] No, you don't have double vision – the viewgraph was really made like *that!*

Alright, my talk by definition won't take as long as the other talks because this work is in progress and not at any final state certainly where we could summarize a lot of results. But it's nevertheless very interesting. Somebody, one of my colleagues at JPL, said "Well, you know, neural networks are nothing more than parallel computational algorithms and architectures, and why don't people call them that?" Well, because I don't think you'd probably get as much money. [LAUGHTER] So anyway this is the hot, new area in signal processing. But I think really the important advances, I believe, are in the applications of this work.

I have a little overview, a definition of what neural networks are [SLIDE #1] from a recent *IEEE Magazine* article by Lippmann [*IEEE ASSP Magazine*, April 1987], a very nice article. If you're interested in or just want a high level overview of what neural networks are and what different types exist, this is an outstanding article.

Basically a network is just an interconnection of all these computational elements, which are called neurons. They all operate in parallel. Do they have to operate synchronously? No, they don't have to be implemented synchronously, they can be implemented asynchronously. There's a lot of work looking at these aspects which is going on at a lot of different places including JPL. Their arrangement in patterns is reminiscent of biological neural nets, which is why they get their name.

What I've done in this viewgraph [SLIDE #1] is I've depicted an example of what a computing neuron might look like. In fact I've shown a little bit more than that. You'll see this again in several of the diagrams that I will present here. I've included here an additional input, the so-called input neurons. These input neurons are special types of neurons which really don't do anything other than just pass the data through. The computing neuron, though, takes the outputs or, in this case, just the input values, the N inputs, weights them by constants or weights them with parameters that can actually be adapted in the course of training this network, called synapse weights, W's. These are added together and the output is thresholded through some type of nonlinearity, and you go on to the next layer to the final output. This would be a mathematical representation of this process. The sum of the weighted inputs go through a nonlinearity and there's typically a threshold, which is some sort of an internal threshold, that can also be adapted. This nonlinearity can be a wide class of nonlinearities, i.e., hard-limiters, signum functions and so on.

I was at a talk a year ago or so and somebody said, "Therefore a neuron is nothing

Application of Neural Networks To Signal Sorting

Edgar Satorius, Ramin Sadr

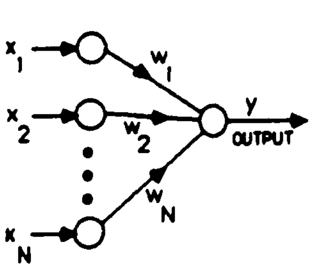
JPL

SLIDE #0

JPL

OVERVIEW OF NEURAL NETWORKS

- DEFINITION: A NEURAL NETWORK IS AN INTERCONNECTION OF COMPUTATIONAL ELEMENTS (NEURONS) OPERATING IN PARALLEL AND ARRANGED IN PATTERNS REMINISCENT OF BIOLOGICAL NEURAL NETS
- COMPUTING NEURON:



$$\longleftrightarrow \quad y = f\left(\sum_{i=1}^{N} w_i x_i - \theta\right) \quad \longleftrightarrow$$

$f(\bullet)$ = nonlinearity (hard limiter, sigmoid, etc.)

θ = internal threshold

- APPLICATIONS:
  - MODELLING BIOLOGICAL SYSTEMS [McCulloch, Pitts; 1943, Grossberg; 1986]
  - SPECIAL-PURPOSE PROCESSORS [Hopfield; 1982-86, Abu-Mostafa, Psaltis; 1987]
  - PATTERN RECOGNITION [Widrow; 1986, Minsky, Papert, 1969]
  - SPEECH ANALYSIS [Gold, 1986; Mueller, Lazzaro, 1986, Sejnowski, Rosenberg; 1986]
  - QUANTIZATION [Tank, Hopfield; 1986, Kohonen, et. al.; 1984, Reed, et. al.; 1987]
  - SIGNAL ANALYSIS [Anderson; 1988, Satorius, Sadr; 1989]

SLIDE #1

more than a summing junction and since the whole world is made up of adders, then the whole world is a neural network." So you could almost write any algorithm like a neural network. Neural networks aren't new by any means. They've been around for a long time and they've been used in lots of applications. Certainly modeling biological systems would be probably one of the first applications of neural networks. They're also used or proposed for doing special purpose processing in general. Abu Mostafa and Dimitras Psaltis and Hopfield, all at Caltech, have done a lot work in this area. They've also been used heavily in pattern recognition.

Speech analysis and recognition is another important area for application of neural nets. If one can develop a nifty way of doing speech recognition then little kids can work computers simply by talking into them. One big technique used in speech recognition is called "Hidden Markov Models". An important competitor to this is a feedforward neural network incorporating back propagation training. So this is a big, hot area in speech analysis. Vector quantization is also an important application area for neural networks as is signal analysis.

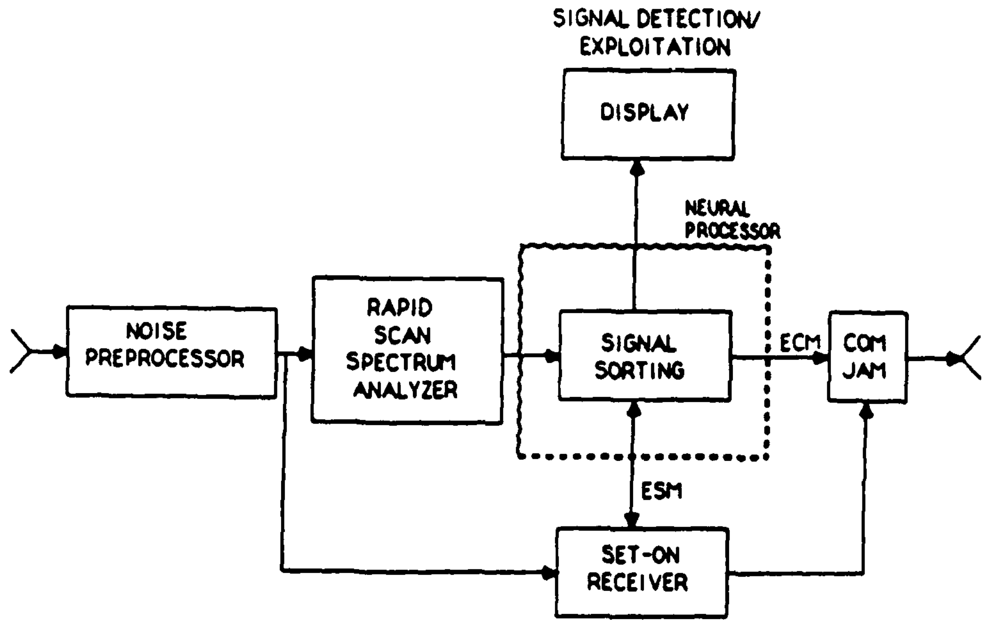
In signal analysis there's probably other work that's going on, but I reference here some work by Jim Anderson at Brown who has been doing some interesting work in applying neural networks to characterizing radar signals. In contrast, we've been looking at communication signal recognition, or identification which is really the first step. The previous talk pointed out that in any identification or recognition problem, there are really two parts: feature extraction and then feature classification. The aspect that we've been looking at initially is feature classification. Here's a picture of a general type of receiver for processing communication signals. [SLIDE #2] You basically have an antenna input and you might have some type of a noise preprocessor so that interference doesn't capture the SYSTEM dynamic range. The major components of these types of receivers are comprised, first of all, of a rapid scanning spectrum analyzer. It does nothing more than scan repeatedly a very broad range encompassing maybe the VHF, UHF-band, HF-band and so on. This rapid scanning spectrum analyzer dumps its data at typically 25 kilohertz resolution so that you may have thousands of frequency cells of amplitude information coming out of the spectrum analyzer in orders of hundreds of microseconds. That's a lot of data that's coming through and can be increased by incorporating multiple receivers so that you can extract additional information, including angle-of-arrival estimates and so on. And this goes into this magic box called a signal sorter. In some applications you may also have a set-on receiver that's connected to the signal sorter so you can extract modulation information. This type of information is used for signal detection, exploitation, or possibly for ECM communication jamming, or just for listening.

The tough problem in building this system is the signal sorter, and the really tough problem is doing it in real time. And this is a tough problem because you're faced with this massive amount of information feeding in from the spectrum analyzer, and you really have to do both the feature extraction and feature identification problems. Well, this is precisely the function that we think might be able to be not replaced but perhaps augmented with, or complemented with, a neural processor.

Now the first sets of experiments that we've been looking at, which I'll go over here briefly,

JPL

GENERALIZED ARCHITECTURE FOR PROCESSING COMM SIGNALS



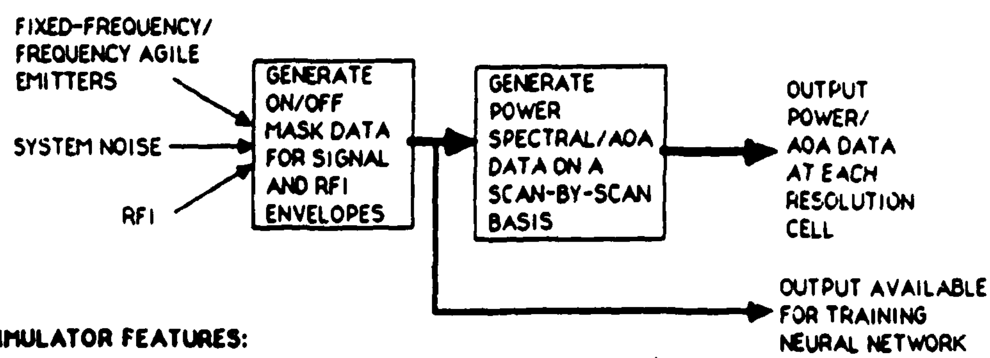
SLIDE #2

California Institute of Technology

JPL

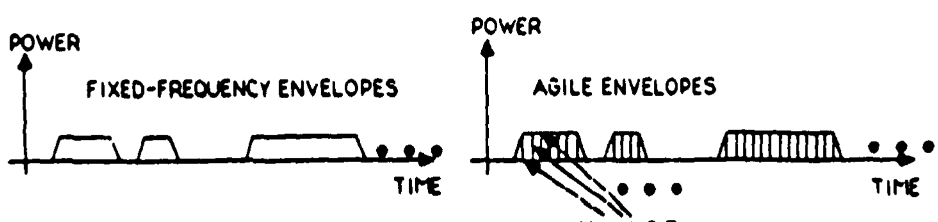
MODELLING AND ALGORITHM DEVELOPMENT

- A SIMPLE COMPUTER SIMULATION PROGRAM (NNSIM) FOR SYNTHESIZING OUTPUTS FROM THE SPECTRUM ANALYZER HAS ALREADY BEEN DEVELOPED:



- SIMULATOR FEATURES:

  - POWER DATA GENERATED BASED ON CHI-SQUARE STATISTICS
  - AOA DATA FOR NOISE-ONLY BINS GENERATED BASED ON A UNIFORM DISTRIBUTION
  - AOA DATA FOR RFI/SIGNAL BINS GENERATED BASED ON NORMAL DISTRIBUTION
  - RFI/SIGNAL ENVELOPES GENERATED RANDOMLY:



SLIDE #3

have been looking at incorporating a neural processor for signal identification. Now there are neural networks and there are neural networks. Neural networks can broadly be categorized into two classes requiring supervised or unsupervised learning. The class that we're looking at uses supervised learning, which means that you've got to train it; i.e., both the synapse weights and internal thresholds are established by training sequences. And you typically train this class of neural networks with data sets in which you somehow have a rich mixture of the different types of the features or the characteristics that you're trying to train this network to exploit. Because once you've trained this network then, until it's retrained, it has to perform signal identification on a class of data which it didn't see from the training. So it has to do a lot of generalization of learning. And that's the basic idea behind the supervised neural networks.

Neural networks based on unsupervised learning can actually extract features from the data, like clustering algorithms, but that's an aspect that we haven't looked at. The field is so rich that we've only been able to concentrate on one part. [SLIDE #3] To do this we set up a simulation model which generated data out from the spectrum analyzer. We also generated angle-of-arrival data – recall that if we did have multiple receivers we could extract angle-of-arrival data. We did this for a couple of signal classes: frequency-agile and fixed-frequency emitters. In addition to simulating the spectrum analyzer output, we also simulate signal mask data (signal on or off) to train the network.

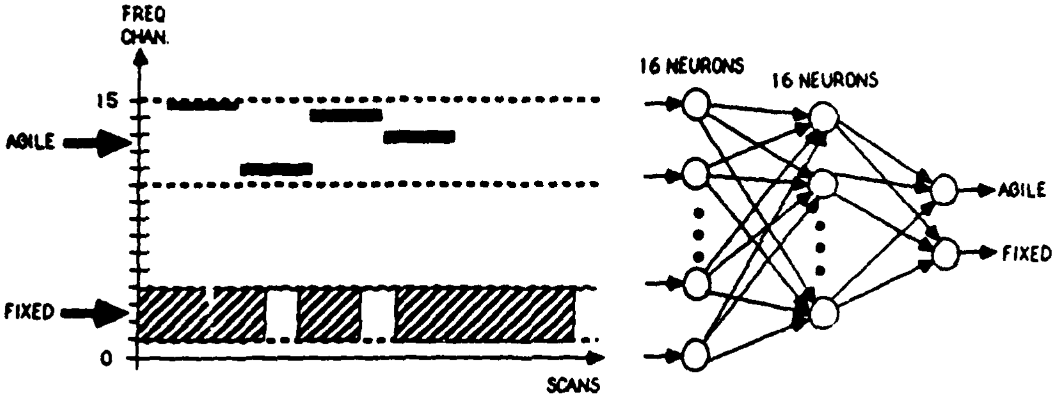
Here's an example of one experiment which we've carried out representing some preliminary results separating frequency-agile and fixed-frequency emitters. [SLIDE #4] In these experiments, we have simulated a 16-channel spectrum analyzer, and in this case the agile and fixed frequency emitters. And all of this spectral information over 16 channels is directly fed into a feedforward multilayer neural network architecture.

What we wanted the network to do is be able to take all this mass of information and separate it, and sort it, and tell us in the end, "Well there was, in fact, during this scan an agile emitter present; during this scan there was a fixed frequency emitter present; both were present, and so on." To do this I can only summarize the result. We did a lot of experiments just to get this far. The experiments were such that we found out that really this problem is a pretty easy one.

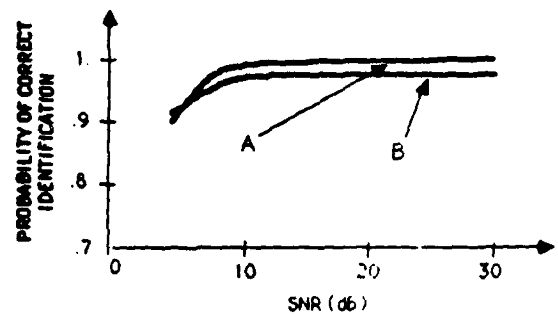
Here's a little bit harder case, but nevertheless an interesting one. [SLIDE #5] This particular case corresponds to a one-channel overlap, fixed-agile emitter class. For this case, we've tried to compare the neural network with a simple energy detector which does not perform a decision when the agile emitter falls into the fixed emitter frequency band. With this type of simple energy detector, you would expect to see a 6% error rate at a high SNR. And that's not quite right because we've got to be a little bit more careful in about how we compute the error rates. But the point is that the neural network does exceed the simple bound (¡ 6% error rate). In other words, the network is doing something a little bit more sophisticated than simple energy thresholding. But what is it doing? It's not exploiting any temporal features of the signal. It can't be, it doesn't incorporate any memory. Based on looking at some of the synapse weights, what we've inferred that it is doing is this: the network started to realize that when the fixed frequency emitter is on, it occupies two bins, but the agile

JPL

SEPARATING FREQUENCY AGILE AND FIXED EMITTERS



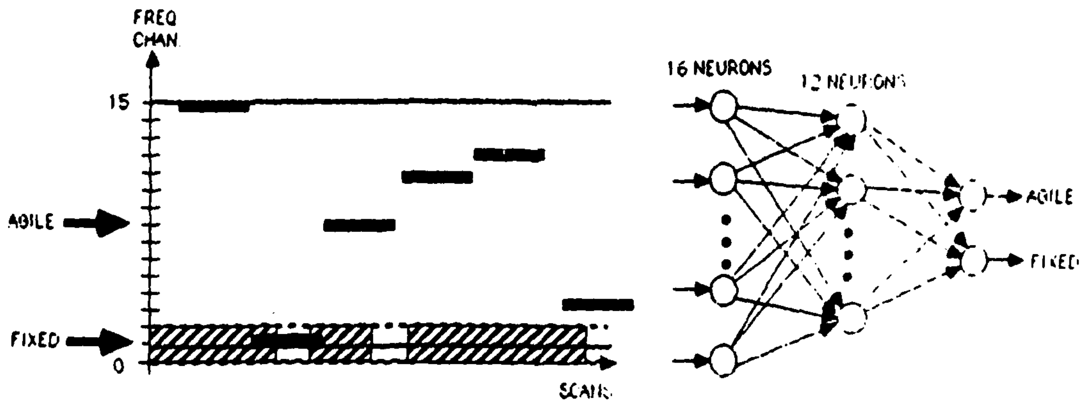
SIMULATION RESULTS
(BASED ON 10^4 TRAINING SCANS)
CASE A: 2 FIXED EMITTERS
CASE B: 1 FIXED & 1 AGILE



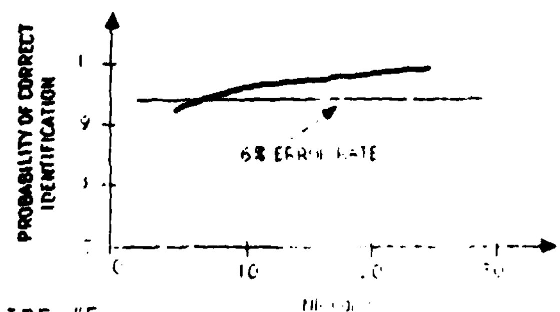
SLIDE #4

JPL

SEPARATING OVERLAPPED FREQUENCY AGILE AND FIXED EMITTERS

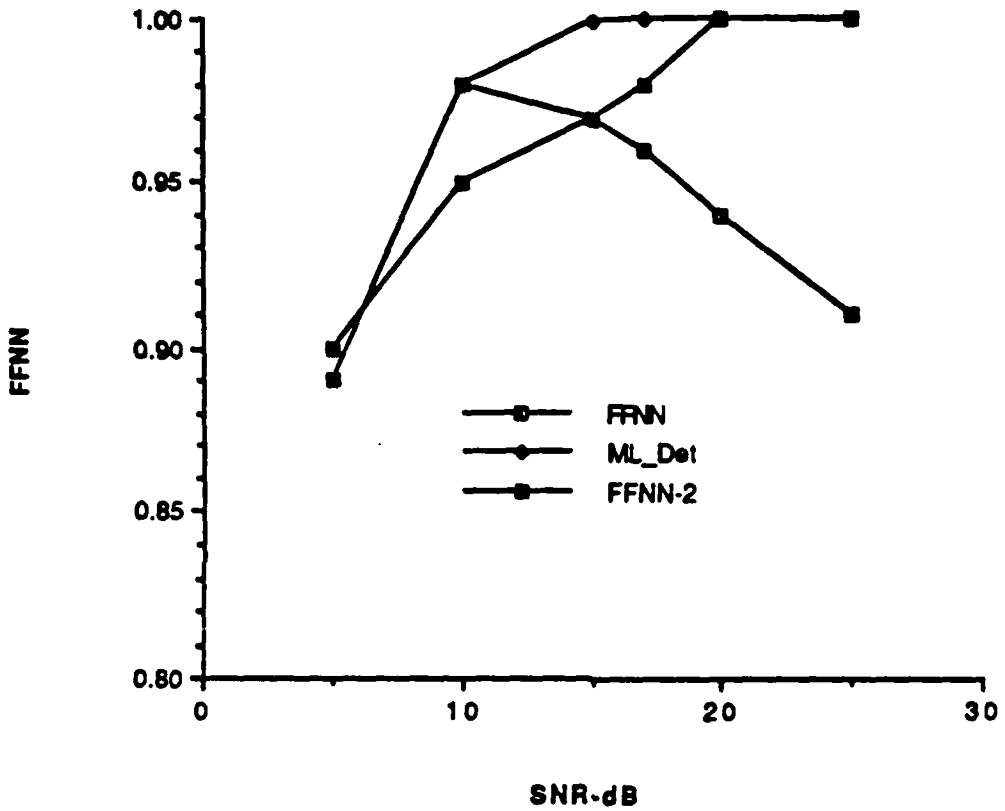


SIMULATION RESULTS
(BASED ON 10^4 TRAINING SCANS)
1 FIXED & 1 AGILE
1 FREQUENCY CHANNEL OVERLAP
- - -> = 6% ERROR RATE FOR
SIMPLE ENERGY DETECTOR @ SNR = ∞



SLIDE #5

Comparison of Maximum Likelihood Detector with FFNN



SLIDE #6

emitter only occupies one. Well somehow it was able to start figuring out and be smart about instantaneous frequency allocations for these different emitters. That's just based on a rough intuitive argument looking at the synapse weights. But something is happening because the probability of correct identification for the neural network classifier is in excess of 98% at high SNR.

Finally, we compared the neural network with a maximum likelihood detector for a simple fixed-fixed non-overlapped emitter scenario. [SLIDE #6] In these experiments, we locally optimized performance of the neural network at each SNR. But at each SNR we always carefully fed noise in and measured the false alarm rate for that SNR. We took the false alarm rates and we used those to compute the thresholds for a dual threshold detector system. As you can see, the maximum likelihood detector outperformed the neural network because it is optimal for this problem. We are currently conducting more comparisons for the overlapped scenarios.

Well, I don't have any more charts, but I just want to leave you with this. Chuck said that one thing he wanted me to end with was where I think the future of neural networks are. Well I don't know where the future of these networks are, because we have so many problems that aren't resolved yet. There's a ton of architectures – we've looked at one. Furthermore, the signal sorting problem is much more complicated than what we've looked at, which is basically identifying an emitter class that we know is present; it's on or it's off. But where is it in frequency? In other words you have to do some sort of tagging of the information. How will that be performed? Via a network? Maybe not. Maybe that'll be something separate and this network will enhance that. What about other information? Angle-of-arrival information? One thing that people say about a network that's very, very neat is that you can simply fuse data types. Well we couldn't simply fuse data types. [LAUGHTER] We tried to simply fuse data types, but the dynamic range problems prevented it. So there's a lot of things that we've got to look at. In other words we have to be able to fuse both angle-of-arrival information as well as amplitude. That's got to help provided the angle-of-arrival information isn't too bad, i.e., isn't degraded too much. What about using multiple scans? Right now we are using a stupid network, because it's not exploiting any temporal features in the data. So what we want to look at is incorporating a buffer of multiple scans up front to see if that doesn't help.

So I don't know what the future is. It's certainly not at the point where I can go off and build hardware, although a lot of people know how to do that. A lot of people know how to build hardware for these feedforward neural networks, but I don't know if many people know really how to fully get this thing operational for a system. You train it forever off-line and then you put it in the field; what if something changes? There are a lot of issues. I don't know where it's going yet, but nevertheless it is interesting.

WEBER: Thank you, Ed. What I'd like to do now is open it up to suggestions and comments. We have two of these mikes that we can pass around. OK ... Ray?

RAYMOND PICKHOLTZ: OK, you hear me?

WEBER: OK, ready ...? Ray Pickholtz.

PICKHOLTZ: My original question which I wrote down was for Mark, but I think it applies also to Bart's talk, and that is: There are many situations, maybe most situations, where you have not one signal present

but where you have many signals present, different signals, different modulation signals, but very similar and very close together. For example you might have multiple signals with the same kind of modulation, with very, very close data rates. It would seem that the proposal that Mark made about using some of these nonlinear processors for extracting rate information would break down when you have multiple signals. And I have a similar question with respect to the kind of classification techniques that Bart talked about. It would seem to me that at the very least you have to do some kind of spatial sorting with arrays in order to isolate individual signals. But even if you do that there's always going to be some remnant of the other signals present, which is going to cause great problems in any of the techniques that I've heard described. I was wondering if we can have some comments on that.

WEBER: Who wants to go first? Bart ...?

RICE: I can go ... can I be heard? Is this working? [TAPPING ON MICROPHONE]

WEBER: Talk right up into it though.

RICE: The co-channel interference problem: Of course, if you can filter things out by direction of arrival or by bandpass filtering then your problem is solved, and you can look at the signals one at a time. But if you have a co-channel interference problem – a long-standing problem – then conventional filtering won't help. Bill Gardner ought to take a mike, because very often you can separate the signals by means of their cyclostationary features. You can either do this in the cyclic spectrum or in the cross ambiguity function (CAF), and Bill has done a lot on showing that different shapes in the cyclic spectrum character the modulation. We've also looked at that, somewhat, and it works; but we haven't gotten anything automatic. Also it depends on how you do the cyclic spectrum. Some of those cyclostationary features are weaker than principal CAF peaks, but let Bill go from here.

WILLIAM GARDNER: Actually I have a comment on a slightly different approach than you were mentioning, Bart. If you can do spatial filtering, in other words if you have an antenna array available, you can in some cases really circumvent the problem of identification of sorting by using blind adaptive algorithms that will automatically select only the signal of interest. For example if you know the baud rate of a signal of interest or, let's say, the chip rate, Brian Agee and Steve Shellen and myself have been studying what we call "spectral coherence restoral algorithms" that will blindly adapt the array to extract only the signal with that baud rate and, if you're within the capability of the antenna array, reject all other signals.

RICE: Well, John Treichler has also shown a constant modulus algorithm will reject certain co-channel interference signals if you start the adaptive weights properly. Maybe he can respond to that. Do you want to, John?

JOHN TREICHLER: I would like to point out one thing, to go back to the original question. The particular problem that Bart was addressing here, that you have the luxury of knowing that any one channel that you're looking at, there's only one signal in it at a time.

RICE: That's right ....

TREICHLER: Now that may not be of broad interest to everybody in this room, but it happens to be the case in the application he was looking at. And that's one of the reasons why the performance is so good, is you never really have to worry about the co-channel problem.

On to the other one, I'm going to hand it right back to Brian Agee in a minute. In fact the early work we did [MICROPHONE MAKING NOISE] ... nice ... the early work we did with the constant modulus algorithm was aimed at correcting multipath, but it was fairly quickly determined that co-channel interference – we would sell it as co-channel, but in fact it needed to be very closely adjacent interference that looked like co-channel – could be nicely rejected by using the very same algorithms. In that case, all we were looking for was simply the property of the signal that, if unencumbered with interference or multipath, it had a constant envelope. And that's really nice except that lots of things have constant envelopes, sinusoids for one; carriers of FDFM signals where the modulation index is quite low appear that way. Lots of things can trick you with that, which was part of the motivation for Brian's and Bill's work on applying more sophisticated ways of deciding if you're happy with the ... I mean, if the signal you're looking for is really the right one. So I'd pass it on to Brian and let him make a comment or two about what you all are doing in that area ....

BRIAN AGEE: I just want to make a short comment that was kind of in the area that Bill Gardner was talking about, and also the comment that Bart Rice made about using the constant modulus algorithm. Some of the recent work that I've been doing in conjunction with some of my clients has been to develop a multitarget constant modulus algorithm. I understand that AST also has developed a multitarget constant modulus algorithm which is capable of sorting, again environments on the basis of constant modulus, and overcomes that co-channel problem. It also can do some identification of environments as whether or not they're signal or interference on the basis of the modulus variation of the signal.

STEARNS: I have something to add. The basic problem of co-channel interference can be solved by attempting to filter the signals in such a way and in such a domain that you do get a single signal present in the analysis band. You can sort the signals either by frequency, that is, determining frequency bands which are sufficiently narrow that you only have a single signal present. You can do it spatially using antenna arrays. You can have algorithms that will adaptively converge to a single signal, and this has been mentioned. A third method is sorting in the time domain. If you're dealing with signals that are not on 100% of the time but that cycle on and off, and if you attempt to make a series of decisions, then you can sometimes find points on which only a single signal is present and the interfering signal is absent. You can look at a long sequence of decisions and look at some metrics on the quality of those decisions, and find times when one signal will appear by itself and then times when the other signal or signals will appear each by themselves, and accomplish the sorting in the time domain as well as in the frequency domain and the spatial domain.

The alternative if you cannot sort the signals out in such a way to have a single signal present, then there's the frontal assault on the problem which is to try to design a hypothesis test that explicitly looks for hypotheses taken two at a time. That is, you'll define a new set of hypotheses which parametrically assume two signals present at a time and parameters for both signals. This, however, is very complicated because the number of parameters is so large that it's not considered a very practical method.

WEBER: It seems like a key issue in all

of this is how much you can get by with assuming you know about the signal up front, period. I mean, if you really don't know very much, if most of your parameters are unknown and you're still trying to look to make a characterization between two signals, essentially the same frequency domain, it could really be tough. Whichever of these algorithms - I don't know if somebody wants to comment on some of that or not ... of the previous commentors.

RICE: Let me make one more comment. There was a recent little vignette that's related in the *Electronic Engineering Times* that was not really explained. If you are faced with colored noise as opposed to an interfering signal, the fellow claims to be able to synthesize the noise by means of fractals. That was in one of the recent issues. He didn't want to tell anybody about it. I think he was holding it as proprietary.

SEYMOUR STEIN: Sort of the Pon-Fleischman ....

RICE: Yes! [LAUGHTER]

STEIN: Got a different question. It's aimed at Mark's talk, but really an open question. You identified power law analysis to get a chip rate line. There are techniques that have been used in modem sync that involve nonlinearities that are not power law. Do you know of anybody who has tried to do detailed analysis of the ability to use those kinds of nonlinearities, you know, for weak signal detection?

WICKERT: We actually did some of that by using a power series model to approximate like a log I_0 function for noncoherent detection.

STEIN: I'm talking about something simpler than that, like taking comp ....

WICKERT: Absolute value, or ...?

STEIN: Taking INQ and just doing absolute I, of course absolute Q.

WICKERT: Not that I'm aware of ....

STEIN: Anybody ....

WICKERT: What I just mentioned is the only thing that we've done. We've looked at that and that was a power series approach, and we could not take it to its limit, but it works. It works just a little bit better if you use matched filtering than if you use like a square law device, so it really doesn't gain that much ....

STEIN: Well what I'm curious about, is that kind of nonlinearity presumably has all powers in it and I'm curious about how it would perform.

WICKERT: Well, that's where our model fell apart because we approached it using a series, and we could only use a finite number of terms. So we were not actually getting convergence.

STEIN: Can I go with a second question?

WEBER: Sure.

STEIN: Also for Mark's talk. You pointed out the tremendous benefits if you want not to be detected of using non-constant envelope signals, you know, by having a filter bite into the waveform. Unfortunately most people who design communications equipment keep insisting on wanting to do efficient amplifiers, which are constant envelope or saturating. In past years there were people who played games with what I tend to call limiter-filter-limiter-filter cascades, trying to generate what looked like signals with interesting properties then limiting them, then trying to put filters on to get rid of the spectral splatter, then limiting again, then filtering again, etc. I'm kind of curious whether anybody that you know of is doing any research into trying to find out how far you might go with the kind .... [TAPE ENDED] ... [LAUGHTER]

WICKERT: The amplitude weightings and pulse shaping approach is a hypothesis. It was just something that just sort of came up after looking at the equations. It's not anything that I have any answers to, it was just a hypothesis.

STEIN: OK, I guess my point is that until somebody invents a miraculous linear amplifier that satisfies comm. designers, it might be well to look at that kind of a problem instead of talking about ... because there are lots and lots of nice ways to generate a Gaussian-looking signal that has very little feature, but they can't be transmitted.

WICKERT: Well, the 4-level signal isn't that complicated, but it's still not constant envelope. So ... I don't know, I don't have anything else to say, I guess, unless somebody else wants to say something.

WEBER: Rob ....

ROBERT PEILE: My question is much more general in nature. I would have thought after this morning's talk that there must be some sort of information theoretic lower limit on the uncertainty over any given set of assumptions on modulation. From these talks it's not clear to me if you're more or less at that limit, a long way from it, getting close, or we're still doubling around. I've got no idea how successful these algorithms are at achieving the theoretical maximum, even under a set of benevolent and unrealistic assumptions.

STEARNS: Well, one way to look at it is that if you're trying to classify communication signals, keep in mind that the communication receiver is able to demodulate these signals at very low bit error rates, looking at intervals of time that are very small. Whereas an intercept receiver, looking at much, much longer intervals of time, perhaps 100 milliseconds, perhaps 2 or 3 seconds, is able to get error probabilities on the order of 10^{-1}. So from that standpoint, the technology that is being used for the algorithms for the intercept receiver are no where near the fundamental bounds.

The key limitation here is complexity. If you look at the type of modeling assumptions that the communication engineer uses when he designs a communication receiver, he does not simply build his transmitter first and then collect signature data on the received waveforms, and then try to design a pattern classifier to decide whether a 1 or a 0 was transmitted. Rather he is able to build a likelihood ratio detector and then put that into the receiver.

On the other hand, the intercept receiver, due to the complexity of building a full model-based algorithm for interception, doesn't have that luxury. So he's forced to go to highly suboptimal techniques in order to identify the signal correctly. One of the things I'd like to see is, as the hardware technology gets better, we can implement more fully the kinds of techniques that we can write out theoretical expressions for, and see those techniques get put into existing systems.

RICE: If you have a complete and accurate characterization of the distribution functions of the features for all of the different signal classes that you are interested in, then you can obtain theoretical expressions for those likelihood ratio functions and you can determine the probability of each decision. Of course, it depends on what the signals are and what the features are. The approach we took was to run through as rich a set of representative test signals as we thought that we would have to handle, whose parameters covered appropriate ranges. Plus, take into account things like the different noise levels and the different channel distortion characteristics. Once you take all that into account, it's

very, very difficult to obtain a valid theoretical expression for the distribution of the features. So, the approach was much more ad hoc.

SATORIUS: You know, I guess that in the results we presented, we tried to do that. But, you know, when you try to do it the case you look at is kind of noninteresting, it's central χ^2 and noncentral.

STEARNS: Once the signal has been mapped into features a considerable amount of information has already been lost. There's far more information about the signal available in the continuous time domain signal than is preserved through the mapping into some finite dimensional feature vector. So from the standpoint of information theory, your key loss occurs right at the very beginning, the very first step in the classification process.

LLOYD WELCH: It strikes me that there's really a sequential process going on. You're talking about sort of the front end. But after awhile, you've identified the modulation type, you have a whole bank of receivers, one of each type that's being manufactured, and you turn on the right one. So there's no reason why you can't do just as well as the communication receiver does. But in line with that succession of steps, I wanted to ask Bart a question. One of the slides that you put up there was the effect of a clustering algorithm that started off with what looked like pretty much of a mess, and wound up with 16 categories and a quadriphase amplitude modulation.

RICE: That's right.

WELCH: I thought that was rather spectacular. Could you tell me a little bit more about that algorithm?

RICE: Yeah, well, that was no more than applying the Godard algorithm, or the constant modulus algorithm, to a QAM signal where we did not know what the constellation was. Once we had despun the carrier, let the constant modulus algorithm converge and recognized the constellation, we switched to the decision-directed mode. The example I showed you had almost no noise on the signal. That's the reason the clusters were so tight.

WELCH: Oh, but it looked like it had a lot of noise to me when ... the first picture ....

RICE: All of that was due to this channel distortion. When the blind equalizer converged a little bit, the channel distortion was removed to the extent that the constellation was recognizable, but not to the extent that the clusters were tight. The tight clusters resulted after going to decision-directed demodulator where we knew what the constellation was.

WELCH: Would that be insensitive to your local carrier being off-phase?

RICE: We had pulled the carrier in by that time. The way that the combined blind equalization technique and carrier despinning algorithm works is to apply the Godard algorithm in conjunction with Jim Mulligan's technique of putting constellation points in each quadrant, and then pretending like that was the actual constellation in despinning the carrier with a relatively small gain, based on its distance from the nearest quadrant center. That method both equalized the channel and despun the carrier. We could be as much as 100 hertz off in our estimate of carrier, and, in blind equalization mode, both equalize and recover and track the carrier.

WEBER: John ....

TREICHLER: I thought I'd amplify on that a little bit. There's been work going on by a number of different people in this partic-

ular area in the last couple of years. I'm not sure exactly who has what and so, competitively speaking, I don't know anything. But I can tell you a couple of benchmarks that I know of for a fact that the place I work, and that is that what you saw here was 16-QAM. Well in fact you can go all the way up to 256-QAM by the same technique quite reliably, with exactly the same algorithm. So as a result the algorithm doesn't need to know whether it's 16-QAM to start with, or 64, or 128, or 256 even. Or that the constellation is a square or a cross or any of the other fancier constellations people come up with now for minimizing the peak-to-average ratio.

Second point is that another way of looking at that despinning that Bart was talking about, we referred to it as "gated decision direction." And you can think of simply gating out any information that's inside of a certain circle in the constellation, leaving only the corners out there. So you simply ... it's just like traditional decision-aided carrier tracking except you ignore any points on the inside, the argument being only the high power points of the edges are the ones you want to trust. And that works quite well so long as the constellation has points. If somebody starts trimming the constellation to where they look circular then you have to play other games, and there are other games you can play.

The third point I want to make, extrapolating a little bit further, I liked Bart's picture with the clusters where he did the clustering algorithm. There are other algorithms you can do also to determine the size of the constellation. Some work we've done at our place and other people have done – the author of the paper sitting next to me, so I shouldn't steal her thunder – but there are other techniques you can use. For example, radon techniques where you simply sort of do averages across the constellation and look and see where the bumps line up. And you can, in situations even noisier than the one Bart showed, you can highly reliably identify constellations up to 256-QAM, even if they're not completely despun. So, for example, say you didn't despin it using this decision-aided technique but simply by measuring the carrier by some other method. For example sometimes you can do a 4th law on the signal and get a little spectral line out there where you think it is 4 times the carrier, and simply go down and try to despin using that, and the constellation isn't quite stopped. Well, some of these other techniques like the radon techniques can in fact work quite reliably even in the face of that sort of uncertainty.

So there are lots of really neat things that are going on for the high order QAM problems, so long as there isn't a tremendous amount of noise or co-channel interference. As soon as you have those problems, you're dead meat. You have to do something else first in order to ....

RICE: I'd make one other comment about the Godard algorithm or the constant modulus algorithm. What that algorithm does is actually adapt filter weights to minimize the dispersion in the constellation points. So those 256-QAMs or 16-QAMs even are far from constant modulus signals. But yet the constant modulus algorithm fixes them right up because what in fact it does is minimize the dispersion in the constellation points, and evidently that's what the channel effect is almost always.

AGEE: I have a question for Bart, possibly also for John. Have you applied your clustering algorithm to the problem of say instead of recognizing the actual signal constellation, recognizing the multiple modulus

that the signal constellation can be on, which would allow it to be used before despinning?

RICE: Are you saying try to assume that there are different modular rings and trying to make a decision as to which one you're closest to? Is that ...?

AGEE: I'm saying trying to recognize what kind of a QAM signal you have by recognizing the different moduli that the signal is on, rather than actually despinning and recognizing the constellation.

RICE: Well, I didn't do that. A fellow at Lockheed-Denver did that a few years ago, but he never really pushed it too far. He didn't do it on any channels that were distorted. He just did it on noisy synthesized signals. He did that on a research effort, and I don't think ever continued to see how well it would work on real channels.

But, yeah, if you do amplitude histogramming, those different moduli, those different QAM types, will have different amplitude histograms. Is that what you mean?

AGEE: Yeah, I was wondering if that had been done.

RICE: Yeah, well he did that and it certainly worked when there was no channel distortion. I didn't trust it until the signal had been equalized somewhat. So, I'm not sure how well that works.

AGEE: I'm just saying if you used a CMA or a Godard algorithm in conjunction with this, it would be able to basically equalize the distortion but it wouldn't be able to despin it unless you used like a Benveniste algorithm, one of those ....

RICE: Oh, yeah, I think that'll work, Brian. I think it will. I mean, I think that my colleague showed that it would work if the channel was equalized. Because the amplitude histogram doesn't take into account carrier offset, you get the same amplitude histogram even if there is a residual carrier.

TREICHLER: This is John Treichler again, and I badly want to give a non-answer. [LAUGHTER] At our own happy little company, there are two rabid camps. One camp thinks this radon transform idea is neat and robust and works all the time and is ....

RICE: Which camp are you in, John? [LAUGHTER]

TREICHLER: There is another camp [LAUGHTER] that believes that this radius method is the right one. The fact that you don't have to have the carrier perfectly extracted, and therefore the constellation perfectly identified, and things like that are very attractive. I have recently heard some of – [UNINTERPRETABLE] euphemistically say this – some highly respected people at your ex-employer who believe that the radius methods are just as good as anything that keeps the two dimensional integrity of the constellation, and they're coming to like that. But the reason I claim this is a non-answer is neither side has analyzed their methods or tested them fully enough to really make any clear claim which is better. It's obvious that if you didn't have to do any final despinning, that would be nice. But some people don't believe that giving up that extra dimension of information is a good idea. So we'll see.

RICE: I might mention that the clustering technique in the picture I showed was not the one that we finally settled on. The one that we wound up using was the following: set a circle at a radius equal to the square root of the energy. Put the hypothesized number of constellation points equally spaced around the circle, and then let them gravitate to cluster centers with an iterative algorithm. If you had the the wrong number, the measure of dispersion would tell you that. If you had the right number, the dispersion measure

would minimize. That idea worked extremely well. There are lots of ways to skin that cat, I think.

WEBER: Andreas ....

ANDREAS POLYDOROS: Thank you. I'm Andreas Polydoros from USC. I'd like to offer a thought or follow on Bob Peile's question ... can you hear me ... on the information theoretic viewpoint of the problem. Let me change a little bit the terminology. I would call it "decision theoretic viewpoint," because you're eventually in a specification problem. You take a decision, a final decision. You'll be surprised to see that if you ask the very simple question, "What is the performance of a joint estimation and decision scheme, the simplest?" The answer is not known. You could answer about the variance, let's say, an estimate by a Cramer-Rao bound. And you can answer the classification probability if you knew the parameters. You ask yourself what is the classification probability of error if I do not know a parameter. The answer is not known. It depends on the nature of the parameter. So if the parameter, for instance, is the unknown data or the phase for which you can write the distribution, you might be able to write down something because of the joint estimator bit correlator of the structure. But if you do not know that parameter and you look at any book or try to write it down, the answer is not known.

So my point is that there are some very fundamental simple questions here that cannot be answered, or have not been answered yet. That even regards a nice additive white Gaussian noise background. You start putting a Middleton model or filtering or so on, you start to realize what the problems become and come about.

Now aside from that I think there's a bigger question that we should be asking, at least from an academic viewpoint. We have been working on this problem at USC also. That is the following: What is the interpretation of all these ad hoc schemes from a likelihood viewpoint, or from a fundamental decision theoretic viewpoint? That is, you do a feature extraction. How does it really relate to a likelihood function? I think an academic job of that sort is very much required because there are so many schemes around, that abound, and you can only start to put them into perspective if you ask yourself that very fundamental question. For instance we've looked at the connection of Bill Gardner's cyclostationary processing to a likelihood function. And he has asked himself questions as to how that connects to ambiguity, processing, or Wigner distributions, and so on. Then, but there still remains the question: How would you connect that to a likelihood functional that takes into account the specific temporal features of the signal and so on? Some of the answers are in. But I think, at least in my mind, there is the dastard world, there's also the academic world, and our job is to push those questions and at least pose them. Some of the things that I see is that we're asking questions: What happens if you have two signals, three signals and so on? But it's important to settle on different classes of models and in each one of them, to try to pose the questions. Because otherwise we're getting confused as to what model applies to where.

WEBER: Bill ....

GARDNER: Yeah, I have some comments in response to Seymour's remarks earlier, having to do with actually both that and the problem of trying to design a signal to be LPI while still not making the communicator's problem too difficult.

If I'm not mistaken I think Mark's analysis

was applied to real PAM signals only. If you consider complex PAM signals which would be the baseband models for the QAM digital signals we'd be interested in, then this idea of, let's say, dithering the amplitude with something like a Gaussian distribution, you would then need a complex value Gaussian distribution, which means you would be dithering the phase as well as the amplitude. But in fact we know that if you dither only the phase you can in fact prevent the regeneration of spectral lines with certain order nonlinearities. But then you're again presented with a problem if you've used a pseudorandom sequence to dither the phase of a digital QAM signal; you may considerably complicate the communicator's problem of undoing that at the receiving end.

If you take another approach of – take the fact that any bauded signal, if it's bandlimited to let's say $1/(N \times \text{the baud rate})$, you cannot regenerate a spectral line with a nonlinearity of order less than N. But if you bandlimited a bauded signal – to such as narrow bandwidth – you've introduced very severe intersymbol interference, again making it very difficult for the communicator. It could be possible, I guess, to reduce the information baud rate by a factor of N, and stuff in between the information bauds known bauds, and perhaps use that to remove the severe intersymbol interference at the receiver. But still I think it seems like you're always stuck with this difficult problem. The more LPI you make the signal, the more difficult you make it for the intended receiver to operate properly.

WEBER: Samir ....

SAMIR SOLIMAN: I have two questions, one addressed to Steve. Do you think a receiver which will approximate the likelihood ratios ... or consider this. Asymptotic behavior of these ratios would be helpful in any way? And the second question, addressed to Ed, do you need to specify a set of features to your neural network receiver?

SATORIUS: Ah, the answer to that question is, yeah, you do. You do that through the effectiveness of the training.

SOLIMAN: OK, in the case of digital likelihood signals, what set of features do you think will be helpful in this case?

SATORIUS: Well, probably several. One would certainly be, which we didn't input yet, angle-of-arrival, one would be amplitude. Another would be basically the timing information that's extracted anyway from the phasing of the hops in the case of the agile emitters. Those would be examples of features that should be put in. It's kind of like the point that I – maybe Steve made about the co-channel interference problem, that timing is helpful in separating out co-channel interference problems, and that would also be helpful for this network.

WEBER: Steve ....

STEARNS: OK, in answer to the first question, the answer is yes. I think it would be helpful to have some receivers that will asymptotically approximate likelihood ratio, or the generalized likelihood ratio. The usefulness would probably depend on whether we're talking about a high signal-to-noise ratio asymptotic receiver, or a low signal-to-noise ratio asymptotic receiver.

I'd also like to comment on Andreas' point, that the performance of the type of structure that I put up is unknown. You'll notice that I didn't present any performance results for it. Not only do we not have the performance, we do not even have bounds on the performance; nor can we obtain them by simulation because of the high degree of architectural parallelism which prevents that struc-

ture from even being simulated on a serial machine or a supercomputer for any reasonable numbers of parameters. The only way to get the performance of that structure is to actually build one, and we're not at the point where that can be done yet.

SATORIUS: But you know that's extremely important from other points of view too. A company builds a machine that's supposed to do this real time identification, they bring it into a sponsor, and the sponsor says, "OK, how well does this work on a very simple problem?" And the guy says, "Oh, well we get an error rate of .0000001 ...," and he can say, "That's neat, you just violated the bounds on any probability you could derive from an information theoretic or from an analysis point of view." So, you know, it's an important issue.

WEBER: With respect to the approximating the likelihood functional at low signal-to-noise ratio, we've been doing some of that with some success on some simple cases. In fact in a couple of cases that corresponded to the spectral correlation algorithm. It ended up being the same function when you simplify the likelihood functional; in most of the other cases it's not. But so far it's had to be relatively simple models in order to really get it to work for reasons that we've been talking about all morning. It's really difficult unless you're really able to assume statistics.

SOLIMAN: Samir Soliman again. I think the idea of using the approximation of the asymptotic type of receiver will help in providing some bounds in the performance of such receivers.

WEBER: That was really one of our motivations for doing that, was to get bounds. Rob ....

PEILE: OK, I'm glad I've put the question of bounds up because it's given me an insight. I get the impression that in 15 years time the performance will be much, much better than these at the moment, which was the fundamental of my question. A second question, which is going back to the QAM question, is that I don't know anyone who uses 256-QAM who doesn't use trellis coded modulation to give you a large Euclidean distance separation over time in an average sense. We've seemed to suggest that if it can help the transmitter and receiver, that ought to be able to help an interception receiver. Has anyone tried to do a multidimensional look at where would this large Euclidean distance help classify coded QAMs? [PAUSE] Maybe the answer's no! [LAUGHTER]

WEBER: Nobody wants to say yes!

RICE: The QAM receiver that produced the tight clusters you saw also had a trellis coding demodulation capability, but we never tried to do anything with trellis coding before we had settled on the constellation and went into decision-directed mode. Once we went into decision-directed mode, then we could crank in the trellis coding but not before.

PEILE: I suppose my point is that the presence of coding over known type ought to help your classification in some sense.

RICE: Yeah, it should. There ought to be a way to do that.

PEILE: Especially Euclidean distance coding.

TREICHLER: I'd like to make a comment on that sometime.

WEBER: John ....

TREICHLER: Ah, John Treichler again. I apologize for once again speaking [LAUGHTER] and I'm going to give you yet another non-answer. Theoretically, yes. Trellis coding can give you another couple of dB. Anybody who has looked at curves of detectability as a function of signal-to-noise ratio where

that couple dB may be crucial. The flip side of it, however, is that it increases the dimensionality of your search. Typically in the intercept column you already have a billion things to look at anyway, and you're often willing to trade down looking for the simplest features first, sequentially, and then only narrowing your search as you proceed through. Typically you'd look for – let's use a case of these newer trellis coded modems that are in, say, voice-band modems. You look first to see if it's 2400 baud. If it's 2400 baud, you try to equalize a little bit and stop the constellation, figure out how big the constellation is. You then go into a decision-directed mode and clean it up enough to where you have fairly accurate bit decisions or symbol decisions. Then you go in and try to identify the trellis, and in fact there are ways to identify the trellis if you know the constraint length, or if you know the constraint length to be a small range of things. And so, yes, you could say I'd like to go all the way out to the end and detect – with no probability distribution whatsoever, everything is equally likely – all trellises, all constellations, all baud rates, all center frequencies. But in fact the dimensionality of that search is so horrible, as a practical matter you don't do it that way.

WEBER: Bill?

STEIN: By the way, I believe if you go back a year or two ago you'll find some literature describing 256 and 1024-QAM modems being used on telephone cable lines. The 256 are not trellis coded, neither were the 1024. The 256, I believe, were indicated to be a real product that was already in operation in the U.S.

RICE: I think MCI uses one ... don't they have a group band modem?

TREICHLER: Yes, there are supergroup band modems and others. An outfit in Atlanta builds one ....

RICE: Super ... I think it's supergroup ....

TREICHLER: ... Digital Communication something or the other is the name of the company.

RICE: I think MCI uses that ....

TREICHLER: And some of them will operate uncoded – yes, they were built for MCI, and they were used to send a T1 over a supergroup. Some of them have forward error correction, not trellis coding; and some of them have no coding at all. What I've heard is that their business has leveled out, then it's expanding no more – 0 into fiber, in the for-what-it's-worth category.

WEBER: One more comment from Bill, and then maybe we'll wind it down ... it's getting time for a break ....

GARDNER: Yeah, I had a comment from a different point of view on the effect of data encoding. Again if you're trying to design a signal to be LPI, you can exploit data encoding. If you take a look at the general spectral correlation properties of all bauded signals, you find that the strength of a spectral line that can be regenerated with nonlinearities depends not only on the pulse shape or envelope, but also on any correlation that might be in the data. So in principle you can introduce correlation through encoding if you reduce the strength of regeneratable spectral lines. Thank you.

WEBER: I'm sure we'll have chances to continue on with this, but thank you for attending. We'll reconvene at 3:00 p.m. with Bill Lindsey's "Communication Channels" session.

INTERACTIVE AND AUTOMATIC SIGNAL ANALYSIS

Bart Rice, Roger Deaton, Clare Heiberger

Lockheed Missiles and Space Company, Inc.

1. INTRODUCTION

Signal analysis is a large and expanding field of activity. There are many reasons for analyzing signals. For example, the classical system identification problem is addressed by analyzing the system output when the system input is known. A flexible modem (MOdulator- DEModulator) which is used to demodulate signals which may have been modulated in one of a number of ways must quickly analyze an input signal to determine the precise type of modulation and parameters (symbol rate, carrier) so that, for example, appropriate tables may be loaded for correct demodulation and decoding. In communications systems, it may be required to route a received signal to an appropriate processor, such as a facsimile machine. These functions have been complicated in recent years by a marked increase in the number of possibilities. An analog channel, for example, may contain a voiced signal or a data signal in one of a variety of modulation formats. The receiving system must quickly determine which processor is appropriate for an incoming signal and allocate resources accordingly.

There are several methods available for performing these analysis functions, either interactively, with an operator or analyst present, or automatically. Interactive signal analysis generally makes extensive use of graphics and may compute measures and features as directed by the analyst. The automatic techniques involve computation of signal "measures" and "features" which characterize the relevant signal types, or at least are sufficient to enable the different signal "classes" to be distinguished from one another on the basis of the computed features.

In the past couple of years, Lockheed Missiles and Space Company has engaged in an Independent Research and Development (IR&D) project which has resulted in systems for both interactive and automatic analysis of communications signals. The interactive analysis system is implemented in software (FORTRAN on VAX/VMS machines, currently with Tektronix 401X graphics terminals, VT100's with Digital Engineering RETROGRAPHICS VT640 cards, or GPX workstations). It integrates a "rule-based" hierarchical decision tree with an internally-developed Signal Analysis and Simulation System (SASS) into a program called RBAS (Rule-Based Analysis System). The analyst, who can specify one of several (currently, two) skill levels at which he or she wishes to operate, is prompted by RBAS to provide answers to a sequence of questions about the signal being analyzed. Based on the answers, RBAS suggests various tests to conduct next (using SASS) and draws conclusions. In its present form, RBAS does not attempt to answer the questions itself or to run the tests automatically. This could be done, however, if an application warranted it.

Lockheed's automatic analysis system accepts time-division multiplexed (TDM) channels sampled at 4000 complex samples per channel per second. In the present configuration, the TDM data are obtained via a digital transmultiplexer operating on a digitized baseband of frequency-division multiplexed channels. The output of the transmultiplexer is placed on a high-speed bus. An internally-developed channel selector board and interface enables selected channels to be passed to either of two Numerix MARS 432 array processors, which are attached to a MicroVax 8200. Algorithms operating in the array processors compute signal features and analyze the signal via a hierarchical decision tree which

uses the feature values to determine the type of each signal.

In section 2, properties of a number of signal types are discussed which enable a signal analyst to identify signals of those types. These are signals likely to be found in analog voice channels. Another IR&D project, not discussed here, has resulted in an Automatic Signal Classification algorithm for signals at Radio Frequencies (RF) using many of same kinds of signal characteristics.

In section 3, Lockheed's Rule-Based Analysis System (RBAS) is described. RBAS is a computer program which incorporates the properties of common voice channel signals into a tool for signal analysts. Rules which guide the analyst based on signal properties are formatted in an almost-English "meta-language." The rules are read by a proprietary FORTRAN program called GENP (GENerate Program), which translates the rules into executable FORTRAN source code. The meta-language and GENP were developed by Tom Wright at Lockheed's Denver Electronics Laboratory.

In section 4, LMSC's automatic voice channel classification system is described. The decision logic employed in the classification algorithm is hierarchical. This approach is compared with a Bayesian approach in which a multivariate distribution function of the features is assumed. Conclusions are summarized in section 5.

2. PROPERTIES OF SOME COMMON VOICE-CHANNEL COMMUNICATIONS SIGNALS

In this section we discuss the properties of various signals which enable them to be quickly identified by a signal analyst. These properties may be incorporated into rules for an interactive "expert signal analyst" or into features for an automatic signal classifier and parameterizer. In the next section, we describe such an expert analyst and automatic classifier operating on multiple, simultaneous input signals.The scope of the discussion is limited to signals which are often used for communications over ordinary telephone channels.

- Voice

Ordinary speech is characterized by rapid energy fluctuations. Figure 1a shows a plot of "long-term energy" (computed over one second) and "short-term energy" (computed over 10 milliseconds). The short-term energy is seen to often cross thresholds above and below the long-term energy value. For most continuous "bauded" signals (discussed below), there are very few threshold crossings. Figure 1b shows a similar plot for a QPSK signal. Thus, the "number of threshold crossings" computed over a specified interval is a feature which helps distinguish voice from other signals.

The spectrum of voice is relatively "spikey", as shown in Figure 2. Also, it tends to change as a function of time much more rapidly then the spectra of other types of signals. Hence, the correlation of the spectrum of voice over different time intervals tends to be lower than that of other signals, providing another useful signal feature. Figure 3 shows the distribution of the correlations of signal spectra over different half-second intervals for different signal types. Note how the values for the bauded signals are near 1.0, indicating that the PSD changes little from one interval to the next.

- FSK (Frequency-Shift Keyed) signals

As in all "bauded" or "keyed" signals, information is transmitted by sending "symbols" at a certain rate, f_b symbols per second, so that the symbol duration $t_b = 1/f_b$ is a constant. In the case of FSK, the symbols are sinusoids at different frequencies. If m symbols are employed, $\log_2 m$ bits of information per symbol may be conveyed. For FSK, usually $m = 2$. If the modulation index (ratio of maximum frequency excursion to symbol rate) is greater than 1, the power spectrum will contain spikes at each of the frequencies corresponding to the symbols. An example is given in Figure 4. The peaks will be narrower if the modulation index is an integer. This is because the sinusoids at either keying frequency are in phase from one symbol to the next. Thus, correlations with a sinusoid (as is a DFT coefficient) at or near one of the keying frequencies reinforce, whereas such correlations may add destructively if the modulation index is not an integer.

Figure 5 shows the spectrum and the spectrum of the square of a Minimum Shift Keyed (MSK) signal, which may be regarded as an FSK with modulation index .5. Computing the spectra of signal powers using the analytic signal rather than the real signal prevents the appearance of spectral energy at certain other frequencies, which is useful since multiple spectral peaks can complicate analysis. When using the analytic signal, the square of an FSK is another FSK with twice the modulation index. Thus, the square of an MSK is an FSK with modulation index 1, so its spectrum exhibits peaks separated by the symbol rate. If the spectra of the signal and its square are presented as in Figures 5, with the spectrum of the square "compressed" by a factor of 2, the peaks in the spectrum of the square "line up" under the actual keying frequencies in the plot of the spectrum of the signal. If the analysis system provides the capability of making screen measurements using a cursor, the analyst can estimate the locations of the peaks in the spectra of the powers and thereby obtain an estimate of the carrier and symbol rate. This technique of inspection of "compressed" spectra of powers of analytic signals is quite helpful in analyzing several types of signals, as will be illustrated below.

Figure 6 shows the result of passing an FSK signal through a frequency discriminator. Figure 7 shows the frequency histogram of such a waveform (an MSK signal, in this case). The output of a frequency discriminator can be degraded by noise. Figure 8a gives an example. It may be possible to restore the two-level nature of the waveform by use of a median filter, which rejects isolated noise spikes (like a low-pass filter) but maintains 'causal' rises or falls in the signal level. Figure 8b shows the result of passing the (oversampled) waveform of Figure 8a through 21-point median filter.

One of the properties of MSK is that its envelope is quite constant, like that of a sinusoid, much more so than amplitude-modulated signals and even more so than phase modulated signals. And, like a sinusoid, its histogram is "u-shaped", indicating that the MSK waveform spends most of its time near ± 1 and relatively little near 0. Figure 9 contains an example.

- MCVFT (Multi-Channel Voice-Frequency Telegraphy)

Frequency-division multiplexing is a well-established method for multiplexing multi-user digital data for simultaneous transmission through a voice channel. CCITT (in English, International Consultative Committee on Telegraphy and Telephony) Recommendations R.31, R.35, R.36, and R.37 specify standards for MCVFT modems. Signals in this general class have very distinctive PSDs, with a "peak-valley" structure which distinguishes them from other signal types. Figure 10 shows the spectrum of an R.31 signal. The individual "sub-channels" or "canals" are on-off keyed (OOK), as can be seen from the "sonogram" of figure 11. R.35, R,36, and R.37 have low-rate FSK modulation in the canals. Some MCVFTs can have low-rate PSK signals in the subchannels.

- PSK (Phase-Shift-Keyed) Signals

These signals are generated by shifting the phase of a constant amplitude carrier at regular time intervals. Rapid phase shifts are wide-band processes, but the signals are band-limited. Thus, energy is lost at the symbol transitions, producing a temporary drop in amplitude of the signal. Consequently, the signal envelope contains a periodic component at the symbol rate. This phenomenon occurs in many bauded signals. It is manifested in the spectrum of the envelope, as illustrated in Figure 12. A symbol-rate component may also be obtained by a delay-and-multiply of the input signal, where the delay is a fraction of the symbol duration. This is the "chip rate detector" commonly used for chip rate recovery in spread spectrum systems.

Let us represent a PSK signal as

$$s(t) = e^{j(2\pi ft + \phi(t))},$$

where the phase $\phi(t)$ "shifts" every T seconds. For BPSK (Binary PSK) signals, $\phi(t) - \phi(t-T) = 0$ or π. Thus, $\phi(t) = \phi_0$ or $\pi + \phi_0$ for some initial phase ϕ_0.

The spectrum of a BPSK (and other PSKs and QAMs) looks like an inverted bowl, while the spectrum of the square of the signal contains a spike at twice the carrier frequency. The spectrum of a BPSK sig-

nal and its square are shown in Figures 13. The reason for this "spectral collapse" is clear from the equation

$$s^2(t) = e^{j(2\pi 2 f t + 2\phi_0)},$$

since $2\phi_0$ is either 0 or 2π, and $s^2(t)$ is a sinusoid at frequency $2f$.

Vestigial sideband signals (VSB) are BPSK or, more generally, ASK (Amplitude-Shift Keyed) signals in which one of the two redundant sidebands has been removed by filtering, except for a "vestige" of the upper or lower sideband. (CCITT Group 2 facsimile signals are also VSB. These "analog fax" signals are easily identified from the End-of-Line dropouts which occur sixty times per second.) The spectrum of BPSK/VSB or ASK/VSB signals is very much like that of an ordinary PSK, but the spectrum of the envelope does not contain a symbol-rate spike. The spectrum of the square will contain a spike at $2f_c$, however. If the signal is basebanded with the recovered carrier, the missing sideband reappears in the real part of the basebanded signal. This is because of the shape of the amplitude response of the VSB filter and the fact that the spectrum of a real signal is conjugate-symmetric. Thus, the spectrum of the envelope (or square) of the real basebanded signal will contain a symbol-rate spike.

For QPSK (Quadrature PSK) signals, $\phi(t) - \phi(t-T) = 0, \pm\pi/2$, or π. Thus, the spectrum of the fourth power of a QPSK signal contains a spike at four times the carrier frequency. The spectra of a QPSK, its square, and fourth power are shown in Figure 14.

Certain PSK modems (MOdulator-DEModulator) will revert to an "idle" state while awaiting data. In this mode, a short repetitive sequence is transmitted. Sometimes the repetitive pattern will be generated by a periodic "randomizer," or "scrambler." In this case, the spectrum will have a line structure, where the lines are separated by 1/(period of the sequence). The relative magnitudes of these spectral lines are obtained by multiplying spikes of equal height by the spectrum of the signal in normal data transmission mode.

If no randomizer is employed, the modem in an idle state may alternate, at the symbol rate, between two phase states. The spectrum of such a signal is likely to have a spectrum which resembles that of an FSK signal. Inspection of the spectra of the powers of the signal often reveal the true nature of the signal. Figure 15 contains the spectrum, the spectrum of the square, and the spectrum of the fourth power of such a non-random QPSK signal. While the first two spectra are atypical of a QPSK, the fourth power reveals considerable spectral collapsing at $4f_c$, which should not occur in an FSK signal.

One form of Differential QPSK (DQPSK) has $\phi(t) - \phi(t-T) = \pm\pi/4$ or $\pm 3\pi/4$. Thus, $\phi(t)$can take on any of eight phase values, $\phi_0 + k\pi/4$, $0 \leq k \leq 7$. But, the spectrum of the fourth power contains two spikes, centered at four times the carrier and separated by twice the symbol rate. The spectra of the first, second, and fourth powers of a DQPSK (CCITT V.26b) is shown in Figure 16.

This phenomenon can be explained by observing that (temporarily assuming that $f_c=0$) $s(t) = s(t-T)e^{j\phi}$, where $\phi = (2k+1)\pi/4$, $-2 \leq k \leq 1$. Thus, $s^4(t) = s^4(t-T)e^{j4\phi} = -s^4(t-T)$. That is, $s^4(t)$ is a signal in "revs" (reversals), with period $2/T$. It is a one-dimensional signal and thus contains strong frequency components at $\pm 2/T$. Reinserting a non-zero carrier f_c, these components appear in the spectrum of $s^4(t)$ at $4f_c \pm 2/T$.

QPSK can be ideally represented by the diagram in Figure 17, where, theoretically, $a(t)$, $b(t) = \pm 1$, and they change to a new value at the same times, every T seconds, where $f_b = 1/T$. Offset QPSK (OQPSK), or Staggered QPSK (SQPSK), is a modulation type in which the keying times for $a(t)$ and $b(t)$ are offset by $T/2$ seconds. The spectrum of an OQPSK signal will look like that of a QPSK signal, and, like a QPSK, the spectrum of the fourth power should contain a spike at $4f_c$. But, the spectrum of the square of an OQPSK will contain spikes at $2f_c \pm .5f_b$ These will be somewhat weaker than those for an MSK signal. MSK and OQPSK are very similar modulations (cf. [4]), despite the fact that OQPSK is regarded as a type of PSK,

whereas MSK is a kind of FSK. MSK may be obtained as an OQPSK with a raised cosine pulse shaping. Often, a demodulator designed for one will also successfully demodulate the other. Also, in the spectrum of the envelope of an OQPSK, the second harmonic of the symbol rate often appears as strongly as the symbol rate itself. This is discussed in more detail in the next section.

8-PSK signals can shift to any of eight different phase states, $\phi(t) - \phi(t-T) = k\pi/4,\ 0 \leq k \leq 7$. The eighth power of an 8-PSK should exhibit a spectral collapse at $8f_c$, but often this spike is quite weak and can be masked by noise. Considerable integration may be required to detect it reliably. Thus, 8-PSK signals are often better identified by other methods. This is especially true in voice-grade channels, where the spectrum of the envelope is likely to reveal a symbol-rate spike at 1600 Hz. Most other PSK (and QAM) signals have keying rates of 1200 or 2400 Hz.

- Quadrature Amplitude Modulated (QAM) signals

The appearance of a symbol-rate spike in the spectrum of the envelope or of the output of a delay-and-multiply, together with a conclusion that the signal is not one of the PSKs, indicates that the signal is a QAM signal. QAMs may be represented as in Figure 17, where $a(t)$ and $b(t)$ may assume multiple values. If we define a QAM signal to be any signal modulated in this way, then PSK signals become a subclass of the general class of QAM signals defined by the equation $a^2(t) + b^2(t) =$ constant.

In theory, for any given QAM modulation, the pairs $(a(t),b(t))$ take on only a small number of possible values. The values form the "constellation" of that QAM type. Some common constellations are shown in Figure 18. A "symbol" consists of one of these pairs, and information is conveyed by the transitions from one symbol to another (in some PSKs, the symbol itself conveys the information, rather than the transition). Symbols are transmitted over specified intervals of duration T, where, as before, $f_b = 1/T$ is the symbol rate. In practice, the pairs $(a(t),b(t))$ take on a continuum of values because of noise, pulse shaping, band-limiting, intersymbol interference, and other distortions due to various channel effects. But, with proper processing, if the noise and distortions are bounded within limits (which vary according to QAM type), the value of $(a(t),b(t))$ when t is near the center of the symbol interval is close enough to the theoretical value of the intended symbol to permit the demodulator to correctly determine which symbol was transmitted during that interval.

As with PSK, for QAM signals the spectrum of the envelope or of the output of a delay-and-multiply process will contain a symbol-rate spike. In general, it should be possible to detect such a spike for any SNR for which it is reasonable to continue the analysis. For example, the delay-and-multiply technique (the classical chip-rate detector) is known to produce a recognizable timing line for BPSKs and QPSKs with negative SNRs. However, for signals with complex, multi-state constellations, the "folding" techniques described above will probably not produce any distinct features which can be used to classify the signals, although the fourth or eighth power of some of these signals may contain detectable energy at $4f_c$ or $8f_c$, if long FFTs and considerable averaging are employed.

Noise presents fundamental limitations on our ability to determine the nature of QAM (or other) signals. Figures 19a and 19b show three-dimensional plots of a CCITT V.29 constellation in which Gaussian clusters are placed about the constellation points to simulate the effect of additive white Gaussian noise at 20 dB SNR and 15 dB SNR, respectively. Even if the precise modulation type of the signal in Figure 19a is identified, the demodulated symbol sequence would contain many errors. And, it would be useless to demodulate the signal of Figure 19b.

Other distortions besides noise, however, have a greater impact on the difficulty of determining the constellation of a multi-state QAM signal. The left half of Figure 20 shows a plot of the pairs $(a(t),b(t))$ at times t in the middle of symbol intervals for a perfectly basebanded QAM signal with no noise added, except that due to (12-bit) quantization. The constellation is unrecog-

nizable because of distortion due to bandlimiting, which produces intersymbol interference. This signal demonstrates a keying spike in the spectrum of its envelope, but the spectra of the second, fourth, and eighth powers contain no identifiable spikes. The signal may be enhanced by adjusting the taps of a transversal filter to minimize some error criterion, under the assumption that so doing will invert the channel response and result in a filter output with much less distortion.

The most common adaptive filter used for this purpose is the Godard Algorithm [5], which minimizes the total amplitude dispersion of the signal. Figure 19b shows the constellation of the signal at the output of a converged Godard equalizer. The Godard algorithm has the advantage that precise carrier information is not necessary for the adaptive filter to converge, whereas other techniques, such as the Benveniste-Goursat [1] method, are more sensitive to error in the estimation of the carrier frequency.

Even for the Godard algorithm, accurate carrier recovery is required to actually recognize the clusters in the constellation. If there is no carrier spike in the spectrum of the fourth power or eighth power, the center of the "inverted bowl" in the signal spectrum may be taken as a rough estimate of the carrier. This estimate may be refined by adjusting the estimate to "track" to an assumed four-phase constellation. This technique is discussed further in the next section.

Once a QAM signal has been sufficiently equalized and basebanded, the signal constellation may be recognized automatically by phase- and amplitude-histogramming, or by clustering. Of course, in interactive analysis mode, it may be identified by visual inspection.

At this point, the QAM demodulator switches to "decision-directed" mode, in which the error signal used in adjusting the phase in the carrier-tracking loop and the taps of the adaptive transversal equalizer is the vector difference between the detected phase and amplitude and the nearest constellation point. Figure 21 shows the results of the demodulation of a (noiseless) 16-QAM when the decision-directed equalizer has converged. All of the original dispersion in the consellations of Figure 20 was due to channel distortion, not noise.

Sometimes the analysis of QAM signals can be facilitated if the modulation is "unbalanced" (i. e., $E(a^2(t)) \neq E(b^2(t))$, where E denotes expected value), as sometimes occurs. The discussion in this case also gives insight into the analysis of offset modulations, such as OQPSK.

As observed above, the envelope of a bauded signal contains useful information. Digitally, this may be obtained by taking the Hilbert Transform of a signal

$$x(t) = a(t)\cos 2\pi f_c t + b(t)\sin 2\pi f_c t$$

to obtain

$$\hat{x}(t) = -a(t)\sin 2\pi f_c t + b(t)\cos 2\pi f_c t,$$

whence $x^2(t) + \hat{x}^2(t) = a^2(t) + b^2(t)$. If, instead, we inspect $x^2(t)$ separately, we obtain

$$\begin{aligned} x^2(t) &= a^2(t)\cos^2(2\pi f_c t) + b^2(t)\sin^2(2\pi f_c t) \\ &\quad + 2a(t)b(t)\cos(2\pi f_c t)\sin(2\pi f_c t) \\ &= .5(a^2(t) + b^2(t)) + .5(a^2(t) \\ &\quad - b^2(t))\cos(4\pi f_c t) \\ &\quad - a(t)b(t)\sin(4\pi f_c t). \end{aligned}$$

The first term is, of course, the envelope squared. The third term has bandwidth approximately equal to the sum of the bandwidths of $a(t)$ and $b(t)$ (assuming they are independent), centered at frequency $2f_c$. It may cause a "hump" in the spectrum of $y^2(t)$, but no spikes.

The second term, however, has the potential for revealing information. Both $a^2(t)$ and $b^2(t)$ contain frequency components at the symbol rate f_b. If $E(a^2(t)) = E(b^2(t))$, then this term contributes nothing to the spectrum. If, however, the modulation is unbalanced, then this spectral component at the symbol rate "leaks through." When it "mixes" with the sinusoid $\cos(4\pi f_c t)$, the result is spectral lines at $2f_c \pm f_b$ in the spectrum of $x^2(t)$. The strength of these spikes is dependent on the amount of imbalance. Thus, estimates of both the symbol rate and the carrier are available from the spectrum of the square.

The second term of the equation provides information also in the case of "offset" or "staggered" modulation types, such as OQPSK, in which symbol transitions in the two "arms" are offset by $T/2$, one-half of a symbol interval. This results in the symbol rate components in $a^2(t)$ and $b^2(t)$ being out of phase by $T/2$ seconds, or 180 degrees. Thus, these components tend to cancel in the first term, $a^2(t)+b^2(t)$, and to reinforce in the second term, again producing spectral lines at $2f_c \pm f_b$. Note that if an offset-modulated signal is basebanded (i. e., $f_c=0$), a strong symbol rate estimate is obtained from computing the spectrum of the difference of the squares of the signal's real and imaginary parts.

The observations above account for one of the principal recognition characteristics of offset modulations. Namely, when the signal is subjected to a non-linear process to produce a spectral symbol-rate line, the spike of the second harmonic is as strong or stronger than the one at the symbol rate itself. Also, intuitively, the symbol transitions in such a signal occur at twice the rate of the keying in each arm, so it is to be expected that there will be a strong spike in the spectrum of envelope at twice the keying rate.

- Partial Response modulations

These modulations introduce deliberate intersymbol interference which can be undone at the receiver, rather than minimize the intersymbol interference, as is done in other modulation schemes. For a good discussion, see [3]. Like VSB signals, these signals do not exhibit a keying spike in the spectrum of the envelope or the output of a delay-and-multiply. Neither do they exhibit a spike at $2f_c$ in the spectrum of the square, as do the VSBs. Class 4 partial response signals are single sideband and usually insert the carrier at a recognizable level at one end of the signal band. The spectrum itself is an inverted half-sine wave, which peaks somewhat more sharply than the spectrum of a PSK or QAM. After basebanding, timing may be recovered by passing the signal through a sequence of rectifications and DC-notch filters and then computing the spectrum, which should contain a keying spike (c.f. [7]).

The keying spike of a Class 1 partial response signal or quadrature partial response signals is at the center of the spectrum. The histogram of a basebanded class 1 or class 4 partial response exhibits a "triangular" structure (Figure 22), the number of causal peaks depending on the number of amplitude states. Also, unlike almost every other type of data communications signal except OOK, partial response signals have a valid state at amplitude 0. Thus, the signal histogram tends to peak at the quantization level corresponding to amplitude 0, even if the carrier estimate is not accurate. Finally, the spectrum of the fourth power of a partial response signal may contain a spike at $4f_c+f_b$, as may be inferred from Mazo [6].

- Summary of signal properties

Many of the signal properties discussed in this section are summarized in Figure 23. Certainly, other types of signals may be found in voice channels. Analog facsimile, various combinations of tones, pulsed signals, and others may also occur. The signal types included here were chosen because they represent a significant portion of the challenges faced by a signal analyst attempting to interpret a sequence of digitized voltage levels emanating from a typical voice channel.

3. RBAS - THE RULE-BASED ANALYSIS SYSTEM

Lockheed Missiles and Space Company, Inc., has developed an interactive "Rule-Based Analysis System" (RBAS) for analyzing and parameterizing the types of signals described in the previous section. Analysis software has been integrated into a configuration-controlled package of VAX-FORTRAN programs called the Signal Analysis and Simulation System (SASS). RBAS may be regarded as a program which gives expert advice on using SASS to analyze a signal. SASS provides an analyst with a considerable graphics capability (the figures in this paper were produced using SASS). It is amenable to implementation in a compact Signal Analysis Workstation, which is currently being developed at Lockheed.

The architecture chosen for RBAS was that of an over-the-shoulder advisor for a

user running SASS. Though RBAS is included in SASS as a subsystem available from the SASS menu, RBAS in its current form does not execute any SASS routines. This is left to the user. RBAS, run on a separate terminal from SASS (or a separate window of the same workstation), prompts the user for answers to various questions about locations of spectral spikes, etc., and makes recommendations about what next to run to classify or parameterize the signal. It also gives instructions on how to execute the SASS programs. Its level of HELP is adjustable, so those users with greater skill get shorter HELP messages. Finally, based on the results of the sequence of results from the various processing steps, RBAS draws conclusions about the signal type.

RBAS is implemented as a set of almost-English rules that are translated by a Lockheed-developed program GENP (GENerate Program) into an executable FORTRAN program. GENP accepts rules in a form such as R14_NOISE = R15_LOW_SNR .AND. R16_FLAT_SPECTRUM and produces FORTRAN code to evaluate each rule. RBAS was developed using GENP rather than some conventional expert system language because of the self-documenting characteristics of the rules, because the rules may be easily changed, or deleted, or rearranged, or supplemented, shortening development costs and increasing flexibility, because the program output by GENP is as fast as FORTRAN, and because it is compatible with the signal processing software routines it may eventually call.

The major task on this IR&D project was the development of rules to classify signals in voice channels. These rules urge the user to look at certain measures of the signal, such as the spectrum, the histogram, the short-term versus long-term energy comparison, etc., and to determine the signal class from the encoded analytic rules. The decision tree used in shown in Figure 24. First, rules to check the quality of the data (SNR, clipping or insufficient use of digitizer dynamic range, etc.) are applied, and the user is given the chance to filter the signal. The signal is then checked to see if it is VOICE by inspecting the graph of short-term versus long-term energy. Of course, a desirable option for the analyst might be to listen to the signal, if a digital-to-analog converter, filter, amplifier, and speaker were available. Such a capability could be easily implemented in software by simply adding additional rules. If a signal is not voiced, spectral, histogram, and other plots are used to further classify the signal as FSK, PSK, QAM, etc. The structure is quite general, so that the ability to handle additional signal types may be included as soon as they are understood and rules are developed to classify and parameterize them. Furthermore, all parts of RBAS interface modularly. For example, if a signal is determined to be QAM, the user has the opportunity to perform QAM parameterization and identification using a program called AUTOPARM (AUTOmatic PARaMeter estimation), a separate RBAS program developed under the same IR&D project.

To aid in the parameterization and identification of a QAM signal, AUTOPARM initially guides the analyst in estimating signal parameters. To obtain a rough estimate of the carrier frequency, several FFTs are computed and averaged to approximate a smooth power spectrum. The carrier frequency is then estimated to be the location of the peak in the smoothed spectrum. This method generally provides an estimate which is accurate to within about .5 percent, which is adequate to allow convergence of the process which follows. Next, a coarse estimation of the symbol rate is made. This is done by finding upper and lower frequencies about the estimated carrier frequency at which the signal PSD is above a selectable threshold. Estimating the symbol rate more precisely is accomplished by computing the spectrum of the envelope of the signal and searching for peaks in the vicinity of the initial, coarse estimate.

Next, AUTOPARM subjects the signal to the Godard blind equalization algorithm and to carrier tracking and symbol-rate tracking loops, to refine the carrier and symbol-rate estimates and to determine the precise QAM constellation. The carrier tracking loop works by adjusting the phase of the received signal to minimize the mean-squared Euclidean distance to a four-phase constellation. It has worked for carrier offsets as large as 100 Hz. Once the

blind equalizer and the tracking loops have converged, a phase amplitude plot readily reveals the constellation, just as in Figure 20.

Detecting the constellation can also be done automatically. Once the equalizer and carrier and symbol-rate trackers have converged, it remains to cluster the detected phase-amplitude points into a constellation. A number of clustering algorithms are available for this purpose. The detected constellation may then be compared with those in a data base (this feature is not implemented at this time). If a match is found, demodulation can proceed using the SASS QAM demodulator in decision-directed mode with the selected constellation and its associated symbols-to-bits conversion scheme. If there is no match, the QAM demodulation algorithm utilizes the recovered constellation and produces a demodulated symbol stream for further analysis.

While the current implementation of RBAS requires inputs and judgements from an analyst, it is clear that these judgements could be made automatically. A peak detection routine would be required to estimate symbol-rate spikes and carrier spikes, for example. In general, GENP is suitable for implementing a hierarchical decision tree and calling routines to provide the answers required to descend the tree. (GENP has been used in precisely this way on another Lockheed IR&D project, on which an automatic signal classification algorithm for RF signals was developed). It should be noted that RBAS is not quite an expert system according to the commonly accepted definition of the term, since it does not provide any traceability features to allow a reconstruction of the path taken through the tree to arrive at the conclusion. But, it marries easily with FORTRAN subroutines and results in correct, efficient code. Incorporation of new rules, modification of existing rules, and rearrangements of the order in which rules are executed are accomplished by editing the rule file and creating a new FORTRAN program using GENP.

The final step in the analysis process is, perhaps, the largest. This is to demodulate the signal to the bit level and analyze the bits. SASS contains such demodulation capabilities, and it also has a substantial bit stream analysis capability, with programs to determine and strip linear randomizers, recover linear recursions, detect known patterns, determine parity check schemes, find frame lengths, and perform other functions which are useful in bitstream analysis. These have not been incorporated in RBAS. At present, the analyst applies RBAS as described above, to help in obtaining the bitstream for further analysis using the Bitstream Analysis Subsystem of SASS.

4. LOCKHEED'S AUTOMATIC SIGNAL ANALYSIS SYSTEM

As part of the Lockheed IR&D program in signal analysis, a system was developed to perform automatic signal classification on multiple simultaneous channels. The signal classification function performed is at a somewhat higher level than that described in the previous section. Signals are grouped into the categories (1) VOICE, (2) FSK, (3) PSK (including PSK, QAM, and Partial Response), (4) MCVFT, (5)NOISE, and (6) UNKNOWN. No parameter estimates are reported in the current version. A block diagram of this signal classification algorithm is given in Figure 25.

The generally accepted approach to automatic signal classification is to compute a set of n signal "features" which "separate" from one another the different signal types to be categorized. These features are usually taken from signal "measures', which are often more costly to compute. For example, the input signal's power spectral density (PSD) estimated using FFT's, may be a (costly) measure, while the "percent f_1- f_2 bandwidth" - the percentage of total signal energy between frequencies f_1 and f_2 - may be a useful signal feature computed from the PSD bins. As a further illustration, Figure 26 shows the distributions by class of a feature which we have termed the "autocorrelation lag ratio" (which was suggested, implemented, and tested by a colleague, George Burgess). It is computed by correlating the PSD coefficients offset by about 30 Hz. Notice how the feature values for MCVFT (R.31, R.35, R.36, R.37) signals are markedly less than those for other bauded signals. An initial feature set was selected by inspecting a large number of

graphic outputs using as inputs signals from each class with a variety of parameters. The signals were subjected to additive white Gaussian noise at several signal-to-noise ratios (SNR's) and to channel filters which provided three levels of amplitude and group-delay distortion. Features whose distributions were overly sensitive to these degradations were rejected.

Features were displayed using one dimensional displays like Figure 26 and two dimensional "scatter plots," as in Figure 27. For each class, three or fewer features were sufficient to separate a given class of signals from the others. The separation using three features f_1, f_2, and f_3 could be displayed with three scatter plots, showing the distributions of each pair of features. When a set of features was chosen, the separation of the signal classes was confirmed statistically using various clustering programs from the IMSL software package. Usually, n, the number of features, should be as small as possible, for two reasons. First, computing the features uses system resources, and second, decisions must be made based on the set of computed signal features. In general, the fewer the features, the simpler the decision logic. This is particularly important when each feature is employed in each decision, as in the case of Bayesian decision logic.

We have opted for a hierarchical decision logic. Paths through the decision tree are taken based on the relationships of features to empirically determined threshold values. Features not needed on a certain path need not be computed, unless deciding whether to compute a feature is as expensive as actually computing it. This can often happen when using array processors for example, since computational tasks may be performed more efficiently than logical ones. The Figure of Merit (FOM) is based on the "closest" decision. If a feature value barely exceeds a threshold value, the FOM is lower than if the feature value clearly exceeds the threshold. Thus, the FOM may be interpreted as an indicator of the confidence on the decision, but it should not be interpreted as a "confidence" in any theoretical probabilistic sense. This choice was made because algorithms which provide a FOM which is a "probability of correct decision" may in fact give the right classification decision but a meaningless FOM. We discuss this further below.

In the Bayesian approach, a likelihood ratio test determines the most likely signal class. Using Bayes' rule,

$$p_i = p(\mathit{class}\ i/\mathit{features}) = \frac{p(\mathit{features}/\mathit{class}\ i)\,p(\mathit{class}\ i)}{\sum_j p(\mathit{features}/\mathit{class}\ j)\,p(\mathit{class}\ j)}$$

where p denotes a probability density function. Thus, if the joint density functions of the features and the *a priori* probabilities of the signal classes are known, then the maximum likelihood decision for the decision class is the one which gives the largest value of p_i. In addition, the value of p_i provides a ready-made FOM. The ratio of the largest p_i to the next largest provides a measure of confidence in the decision. Usually, for computational reasons, log-probabilities are used, so that multiplications and divisions are replaced by additions and subtractions. If Gaussian probability density functions are assumed, the log-probability ratios may be computed without taking exponentials. Also, if feature values computed over disjoint time intervals may be regarded as independent, then the log-probabilites $\log p_i$ may be added to form an accumulated FOM. Thus, if the evidence in favor of any class is insufficient (i.e., $\log p_i$ is too small in relation to the nearest competitor), the decision may be deferred and another set of features computed.

As attractive as this approach appears at first, there are several difficulties. First, the *a priori* probabilities must be known. Usually, they are assumed to be equal, but this assumption is probably incorrect for most real environments. These can be refined by accumulating channel statistics, and the Lockheed system has that capability. But, this means that the rarer, more interesting signal types, may be less likely to be declared when they occur. This difficulty is eliminated in the hierarchical approach by attempting to classify the most likely signals first. Precise *a priori* probabilities are not necessary, only a rough order of frequency of occurrence and/or priority.

Second, the distributions $p(features/class=i)$ must be known for each class. They can be computed using empirical data and stored, but this requires a considerable amount of storage, and care must be taken or it can also be unreliable. Slight variations in computed feature values from those used to "train" the system can result in the features falling in sparsely populated bins, thereby producing low values of $\log p_i$ for the correct class. Usually, the distribution is assumed to be multivariate Gaussian and parameterized. This assumption can be quite incorrect, depending on the features employed, and can lead to meaningless parameters for the distribution. By simply comparing feature values to thresholds as in the hierarchical approach, algorithm performance is relatively insensitive to the precise form of the distribution.

Third, as mentioned above, it is desirable to minimize the number of features. This can be done statistically from a larger feature set (cf. [2]), but then the reduced set of features may lose their "heuristic" value in that it may no longer be possible to interpret a given feature as capturing some salient characteristic of some signal class. If a new signal class is added, the reduced feature set may or may not provide adequate separation of the signal classes. The chances that it will are less for a "reduced" feature set. By using a larger feature set as in the hierarchical algorithm, there is a greater likelihood that a new signal class may be accommodated without modifications to the feature set, but rather by additions to the decision logic.

The fourth disadvantage is that all features must be computed to compute $\log p_i$, even if the signal being analyzed may be categorized on the basis of only a few features. This may not be serious, since the principal computational cost is incurred in obtaining the measures, as mentioned above, (e.g. FFT or instantaneous frequency) not the features (e.g., ratio of sums of some FFT bins). However, if unnecessary measures are computed, then the extra computation can be significant. This problem can be avoided in a hierarchical algorithm, since features (and measures) need only be computed as needed, as mentioned earlier. This advantage can be insignificant or critical, depending on the application.

The fifth and perhaps the most subtle of the difficulties with the Bayesian approach is that the distribution of feature values for an incoming signal may be different from those of the overall class to which that signal belongs. In particular, the assumption that the feature values are independent over disjoint time windows may be accurate for the signal itself but highly inaccurate for the overall class to which the signal belongs. Again, the hierarchical approach is affected little by perturbations in the feature distributions such as this. An additional advantage to the hierarchical approach is that it can be implemented using GENP. This was not done in the system under discussion here, but, as mentioned earlier, it was done on another IR&D project which developed an automatic classification program for RF signals. In that algorithm, incidentally, the ability to compute features only when needed provided a very significant computational advantage.

Still, in spite of these objections, a Bayesian classifier can produce excellent results (i.e., accurate classification decisions). And, it has the huge advantage that the system can be easily "trained" for each signal class to obtain the distributions needed to compute the values $\log p_i$. The choice of which approach to take - Bayesian or hierarchical - ultimately depends on the particular application.

- Lockheed's Voice Channel Processing Module (VCPM)

Lockheed's automatic voice channel signal analysis system consists of six major components (see Figure 28) : a Frequency-Division Multiplexed (FDM) demultiplexer, a voice-grade-channel distribution (VGCD) system, a continuous channel processor (CCP), classification algorithms, a card cage controller (CCC), and a station controller (SC). The FDM demultiplexer is a "transmultiplexer" which separates input FDM basebands into individual four-kilohertz channels an re-multiplexed into TDM (Time-Division Multiplexed) channels. The TDM channels are passed to the VGCD system, which puts them on a high-speed distribution bus, to make the data available to

other processing units. The CCP selects voice channels from the high-speed distribution bus and acts as a front-end processor to the classification algorithms. The classification algorithms receive the channel data from the CCP and perform a "coarse" classification as to whether the signal is voice, noise, or some other type of communication signal. They perform a more detailed "fine" signal classification on a selected subset of the channels. Classification decisions, FOMs, and other information are passed to the station controller, which logs the data to disk files for post processing and to displays for real-time analysis.

The FDM demultiplexer, the VGCD system, and the CCP are implemented in hardware and reside in two card cages. The hardware is controlled over a high-speed VME bus by a CCC. The algorithms are implemented in software which resides in two high-speed array processors. A given channel can be selected for fine classification in either processor. High-level hardware and algorithm control is obtained through a station controller, which is implemented on a general purpose computer.

- FDM Demultiplexer

The FDM demultiplexer demultiplexes 150 channels from a 150-channel input FDM baseband or 300 channels from either a 300 or 600 channel input FDM baseband. The 300 channels to demultiplex within the 600 channel baseband are selected through a digital sub-band tuner within the demultiplexer. The FDM demultiplexer accounts for differences in sampling rates by interpolating the input samples to an internal sampling rate. This flexibility within the FDM demultiplexer is accomplished by downloading parameters from the controller.

- Voice Grade Channel Distribution System

The voice grade channel distribution (VGCD) system accepts output from up to eight demultiplexers and time division multiplexes the data onto a high speed distibution bus. Only one FDM demultiplexer is currently implemented. The VGCD controls three diagnostics for on-line status monitoring. In the first diagnostic, a test pattern generator unit outputs a test tone which is added to the FDM baseband at the input to the FDM demultiplexer in a channel known to be a guard band or an inactive channel. The input baseband is demultiplexed, and the test tone data is passed through the system and monitored by the algorithms. This test serves as an overall system health test. In the second diagnostic, psuedo-random bit patterns are taken from the test pattern generator when data are unavailable from a demultiplexer. The patterns are passed through the TDM multiplexer units and monitored by a test pattern receiver unit. The test pattern receiver unit monitors the test data on a bit-by-bit comparison basis and flags errors to the controller. In addition to these two tests, each unit within the VGCD generates parity bits and checks blocks of data to see that the appropriate parity checks are satisfied.

- Continuous Channel Processor Hardware

The continuous channel processor (CCP) hardware selects the 300 FDM-demultiplexed channels from the VGCD bus. It then heterodynes the channels upward in frequency by 2000 Hz from their initial frequency band between -2000 and +2000 Hz to the band 0 - 4000 Hz. The data are then transformed using a 40-point Winograd-Fourier Transform (WFT), implemented in hardware. An integer-to-floating point conversion is performed, and the CCP then passes both the temporal and spectral data to the array processors for signal classification.

- Classification Algorithms

The classification algorithms in the array processors can be broken into two distinct algorithms : the coarse classifier (CC) and the fine classifier (FC). The algorithms are controlled within the array processor by algorithm control software which accumulates the algorithm decisions and related information and communicates the decisions to the station controller.

The CC is the first algorithm to examine the spectral channel data received from the continuous channel processor and is responsible for activity recognition, high-level classification of the signal type, and signal-to-noise ratio (SNR) estimation on 300 channels. The CC algorithm performs

classification of signals as VOICE, NOISE, or OTHER and returns its decision to the station controller every ten seconds. Along with the classification decision, the algorithm returns a classification figure-of-merit (FOM) and an SNR estimate accurate to ± 1 dB. Consecutive noise decisions are used to declare channel inactivity.

The fine classifier (FC) examines up to fifty channels of temporal data received from the continuous channel hardware and separates input signals into six categories : VOICE, FSK, MCVFT, PSK/QAM/PR, NOISE, or UNKNOWN. The fine classifier returns the classification decision and a classification FOM to the station controller every 10 seconds.

- System Control

The station controller (SC) functions as the operator interface to the VCPM system. The station controller is responsible a number of functions, including high level control and status of the hardware and algorithms, display of algorithm classification decisions on tasked channels, maintenance of system tasking logs and algorithm decision logs, and generation of processing history and analysis reports.

VCPM system control is accomplished principally through a VAXstation 100 attached to a VAX 8200. Three other display devices allow the operator to monitor system status : the hardware status display, the algorithm status display, and the algorithm graphics display. The hardware status display is a SUN 2/160 color monitor attached to the card cage controller. The algorithm status display is a VT220 terminal attached to the VAX 8200, and the algorithm graphics display is a color monitor attached to a second SUN 2/160. A future version of the system has been planned in which all of these functions will be available in single display device.

The card cage controller (CCC) is responsible for low-level control and status of the FDM demultiplexer, the voice grade channel distribution, the T1 trunk interface, and the continuous channel processor hardware. The CCC downloads configuration parameters to the hardware in response to configuration tasking commands received from the station controller. The CCC also monitors the hardware for error conditions and flags the error conditions to the station controller as detected.

5. CONCLUSIONS

The vast majority of signals in voice channels may be automatically classified by computing signal "features" which capture salient characteristics of the possible signal types and by performing some form of cluster analysis on the detected features. Lockheed Missiles and Space Company, Inc. has developed a system for automatic analysis of multiple input signals using these principles.

Interactive signal analysis involves use of a combination of visual graphic aids and parameter measurements. It can be facilitated by Lockheed's Rule-Based Analysis System (RBAS), which guides an analyst through the analysis of an input signal by means of a sequence of questions and conclusions drawn from the answers to those questions. Such a signal analysis "expert" can operate at one of several skill levels, with the system providing more or less help to the analyst, as needed. Varying levels of automation to the analysis process are possible also.

5. References

[1] Albert Benveniste and Maurice Goursat, "Blind Equalizers," *IEEE Transactions on Communications,* Vol. COM-32, No. 8, August 1984, pp. 871-883.

[2] Richard O. Duda and Peter E. Hart, *Pattern Classification and Scene Analysis,* John Wiley and Sons, Inc., New York, 1973, pp. 429-458.

[3] Kamilo Feher, *Advanced Digital Communications,* Prentice Hall, Inc., Englewood Cliffs, N. J., 1973, pp. 114-121.

[4] K. Feher and B. Morais, "Bandwidth Efficiency and Probability of Error Performance of MSK and Offset QPSK Systems," *IEEE Transactions on Communications,* Vol. COM-27, No. 12, December 1979, pp. 1794-1801.

[5] Dominique N. Godard, "Self-Recovering Equalization and Carrier Tracking in Two-Dimensional Data Communications Systems," *IEEE Transactions on Communications,* Vol. COM-28, No. 11, November 1980, pp. 1867-1875.

[6] J. E. Mazo, "Jitter Comparison of Tones Generated by Squaring and by Fourth Power Circuits," *The Bell System Technical Journal,* Vol. 57, No. 5, May-June 1978, pp. 1189-1199.

[7] J. D. Olson, "A Technique of PCM Transmission over 2 GHz Microwave Radio," *IEEE Transactions on Communications,* Vol. COM-23, No. 9, September 1975, pp. 1004-1007.

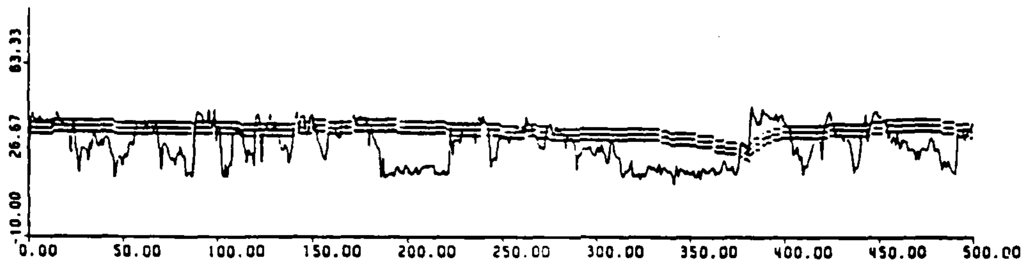


Figure 1a. Short-term versus long-term energy - voice

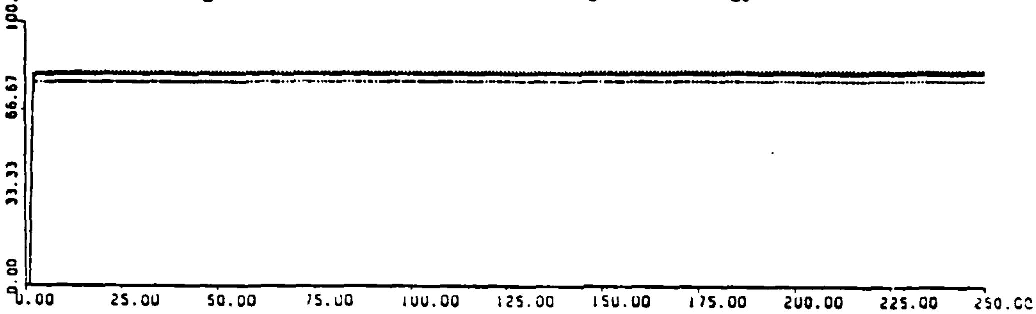


Figure 1b. Short-term versus long-term energy - QPSK

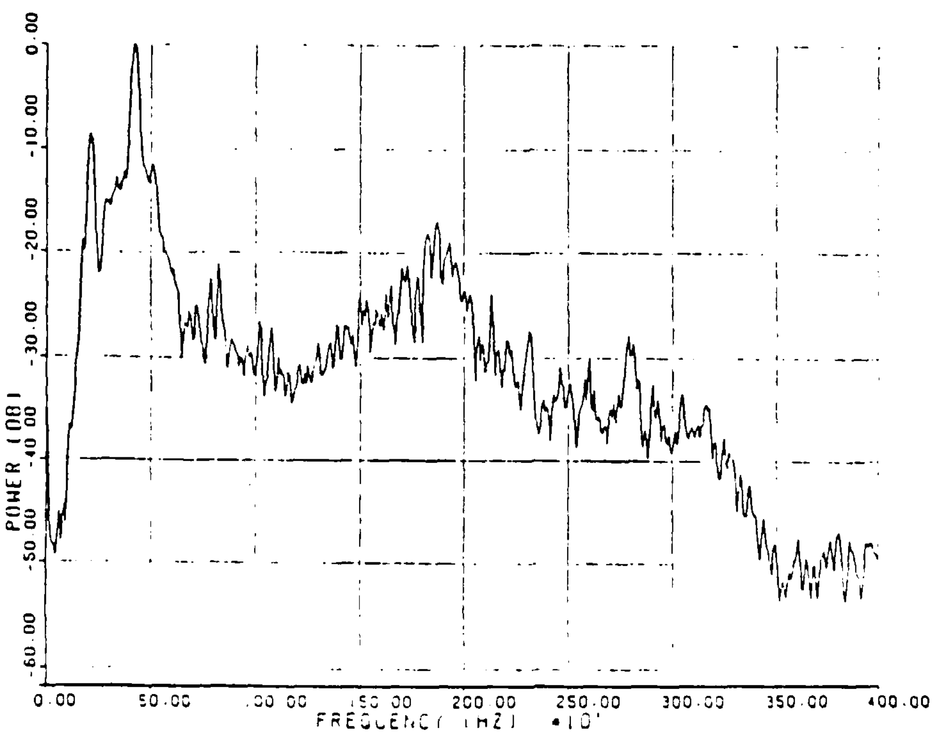


Figure 2. Power spectrum of a voice signal.

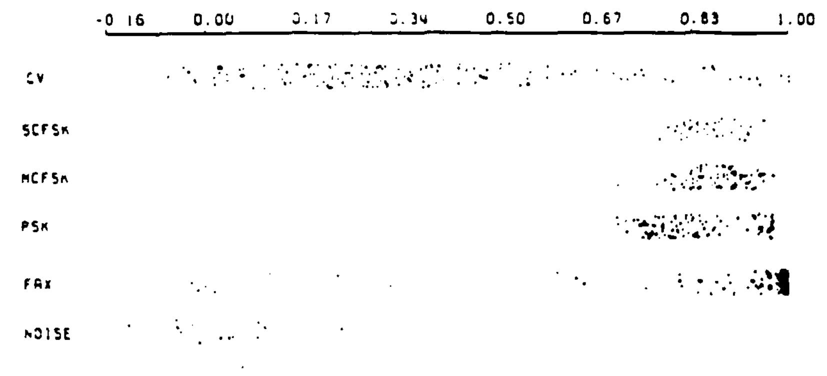


Figure 3. Correlation of spectra over independent time intervals.

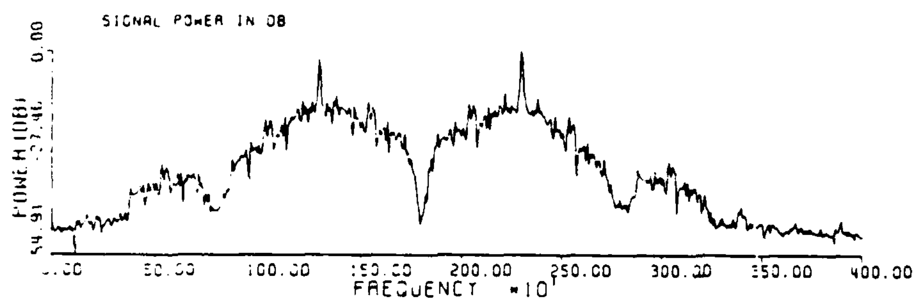


Figure 4. Power spectrum of FSK with integer modulation index.

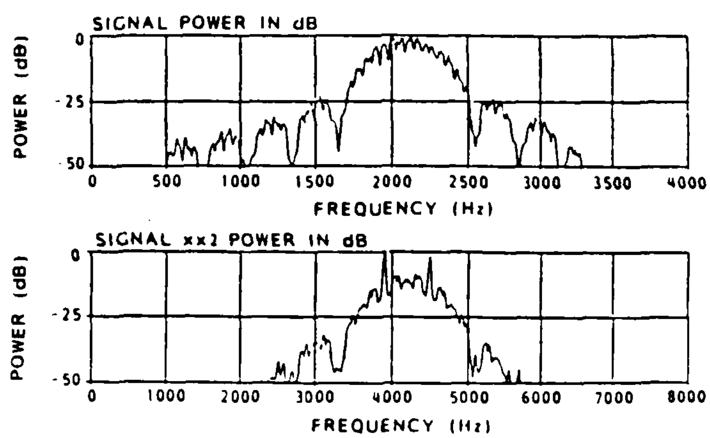


Figure 5. Spectra of MSK signal and its square.

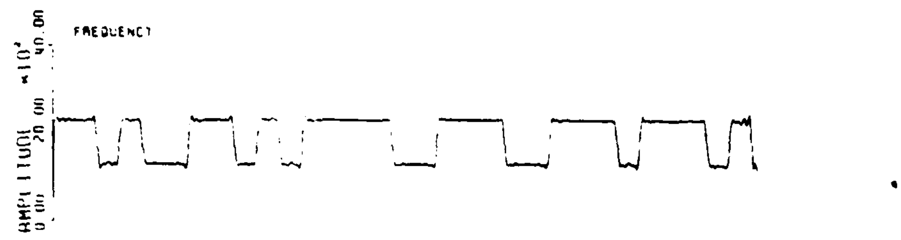


Figure 6. FM-discriminated FSK signal.

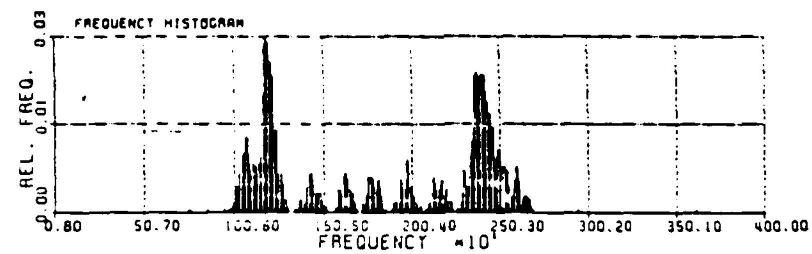


Figure 7. Frequency histogram of MSK signal.

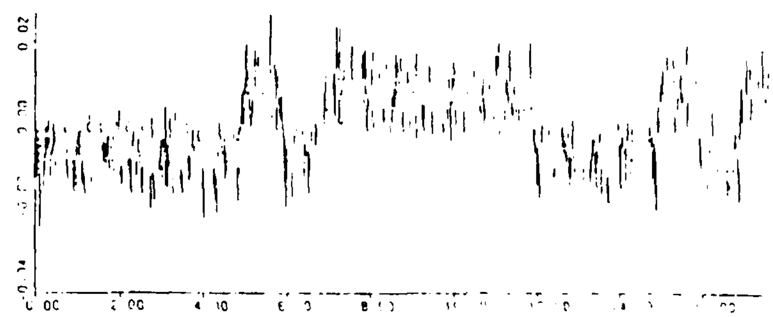


Figure 8a. Demodulated FSK. SNR = 10.7 dB.

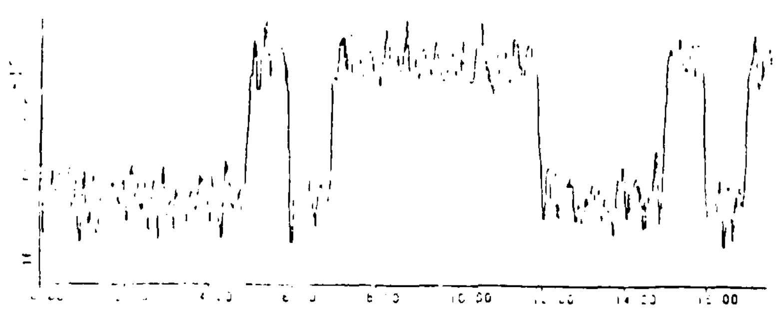


Figure 8b. Waveform of 8a after applying 21-point median filter.

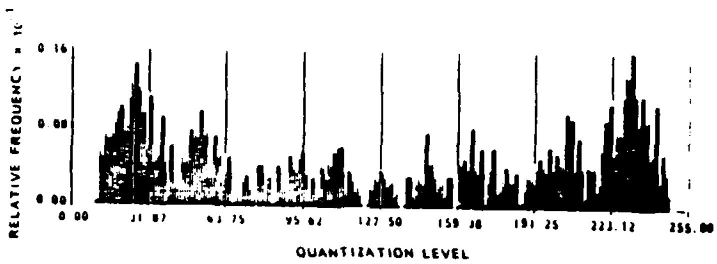


Figure 9. U-shaped histogram of MSK signal.

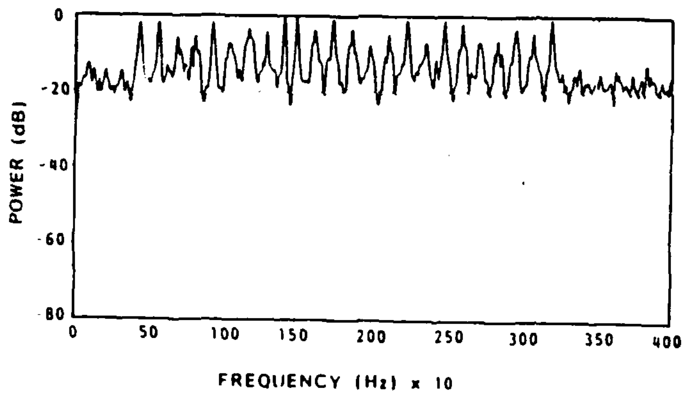


Figure 10. Spectrum of R.31 MCVFT.

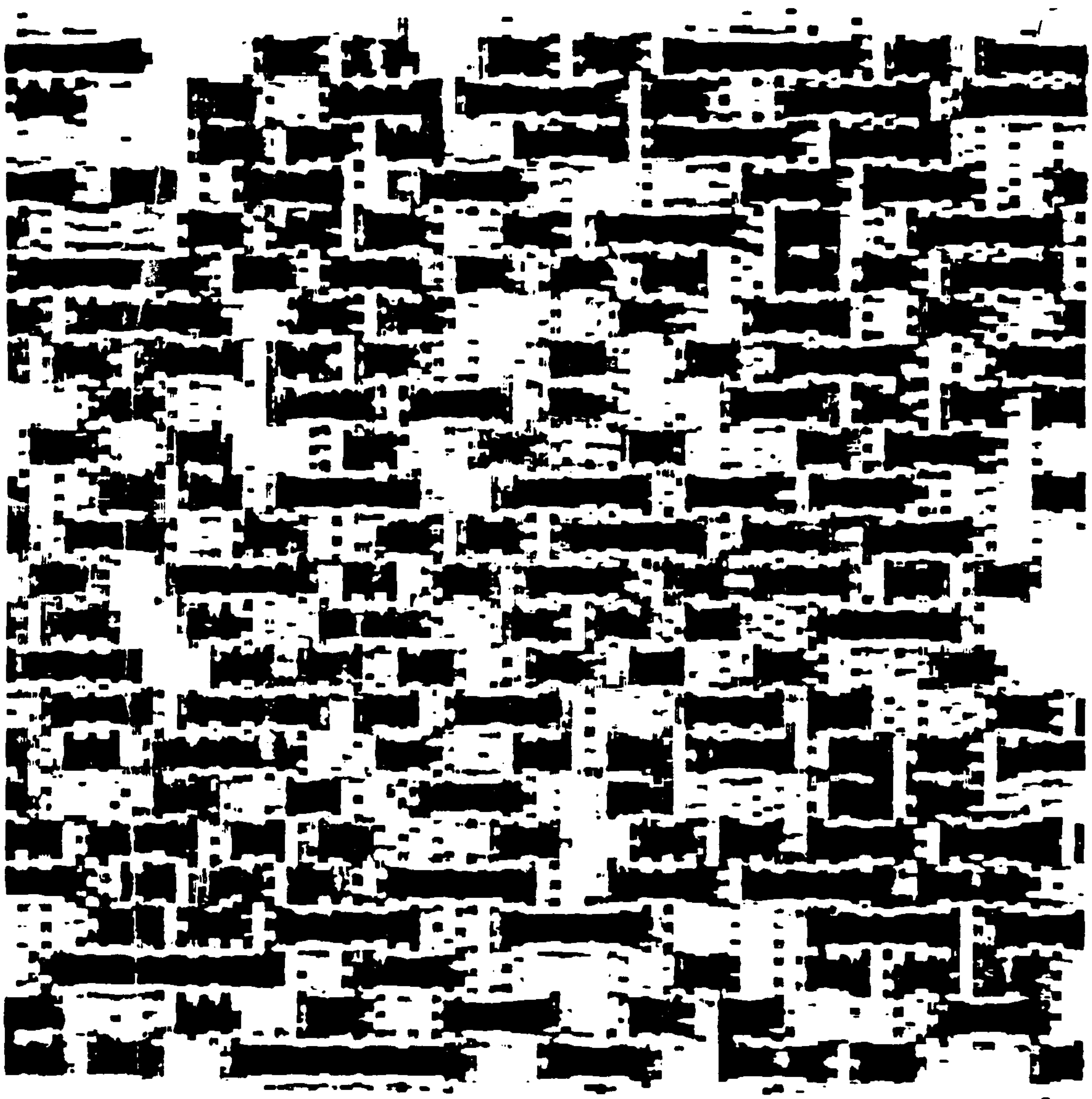


Figure 11. Sonogram of R.31, showing the 24 50-baud OOK telegraphy subchannels.

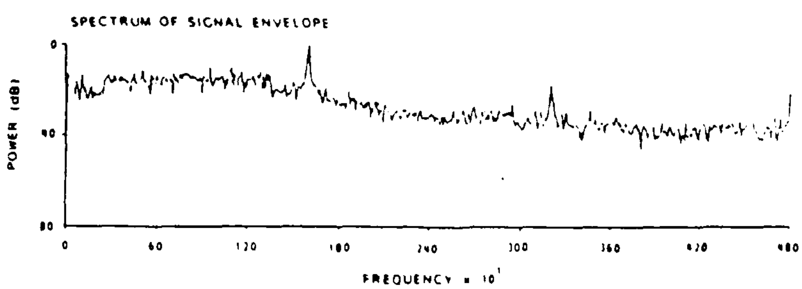


Figure 12. Spectrum of envelope of PSK signal.

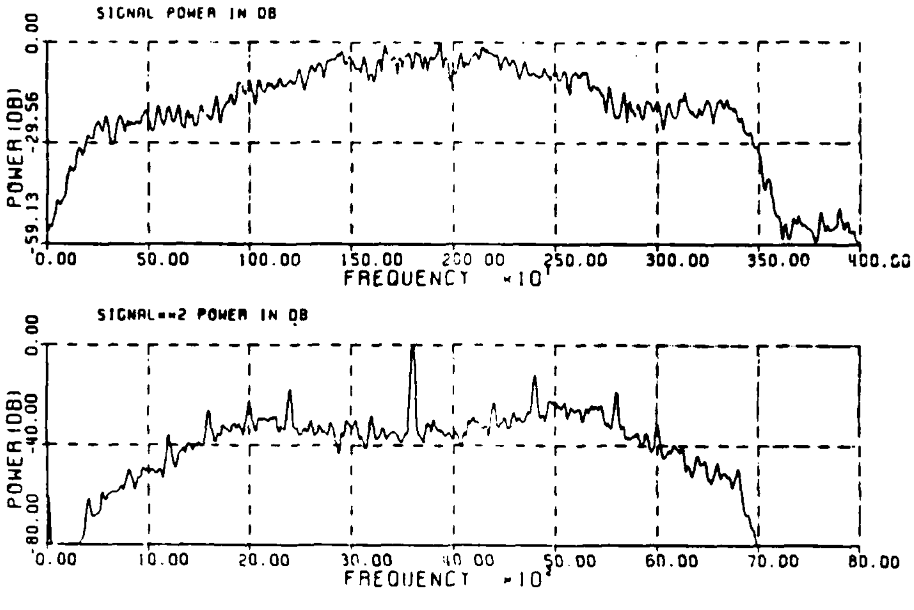


Figure 13. Spectra of BPSK signal and its square.

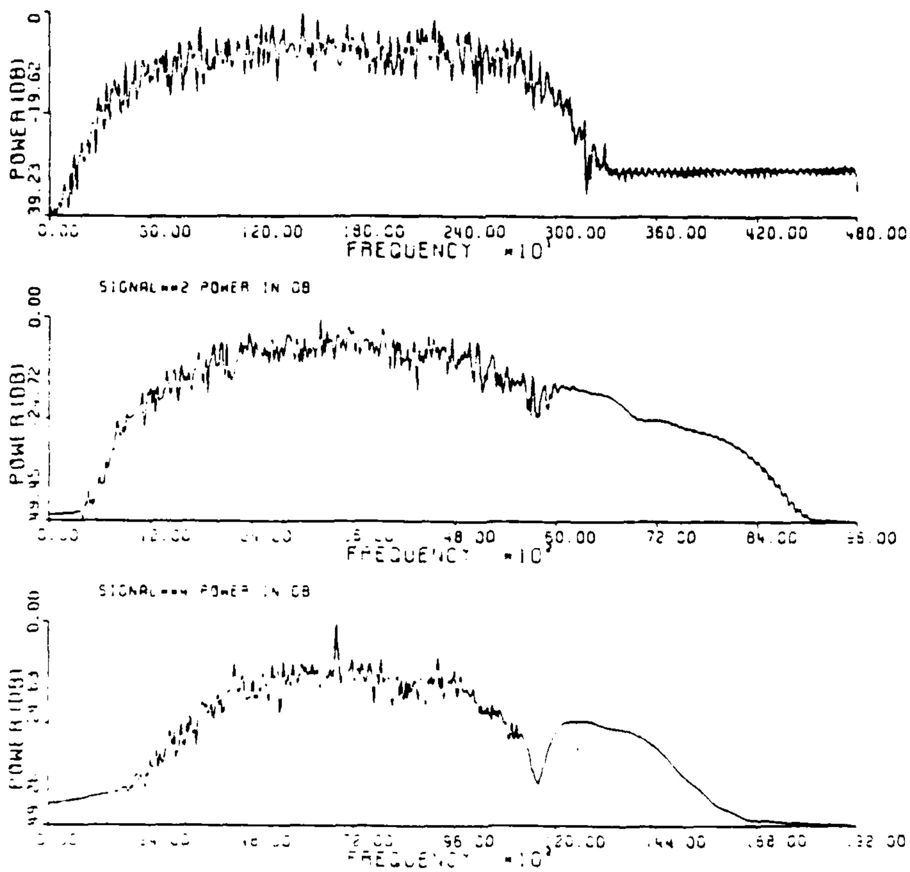


Figure 14. Spectra of QPSK signal, its square, and its fourth power.

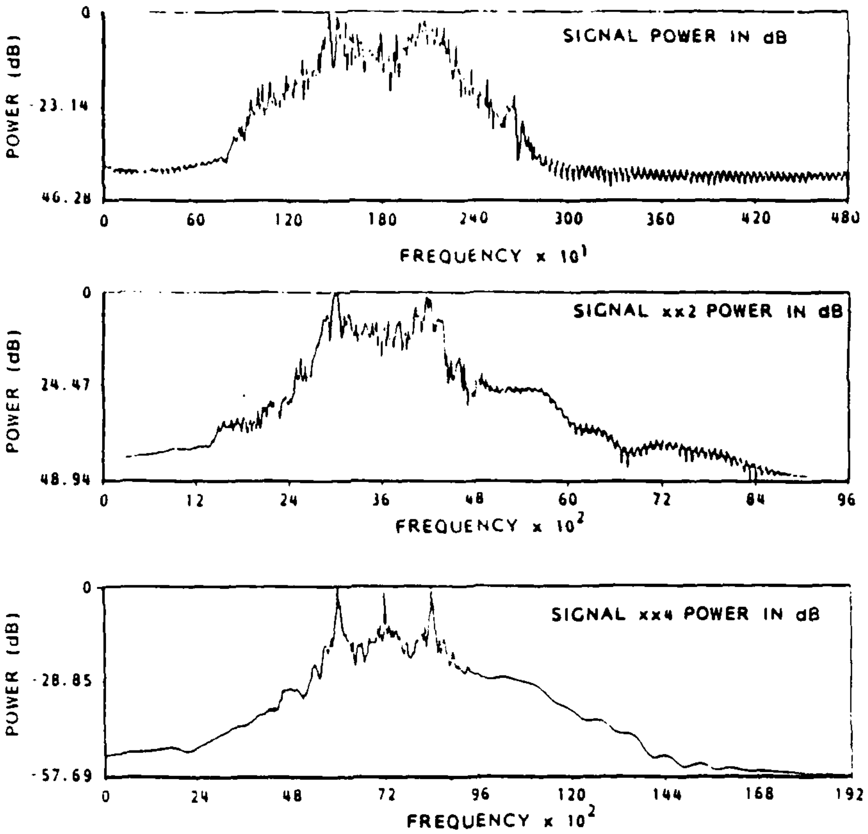


Figure 15. Spectra of powers of a non-random QPSK.

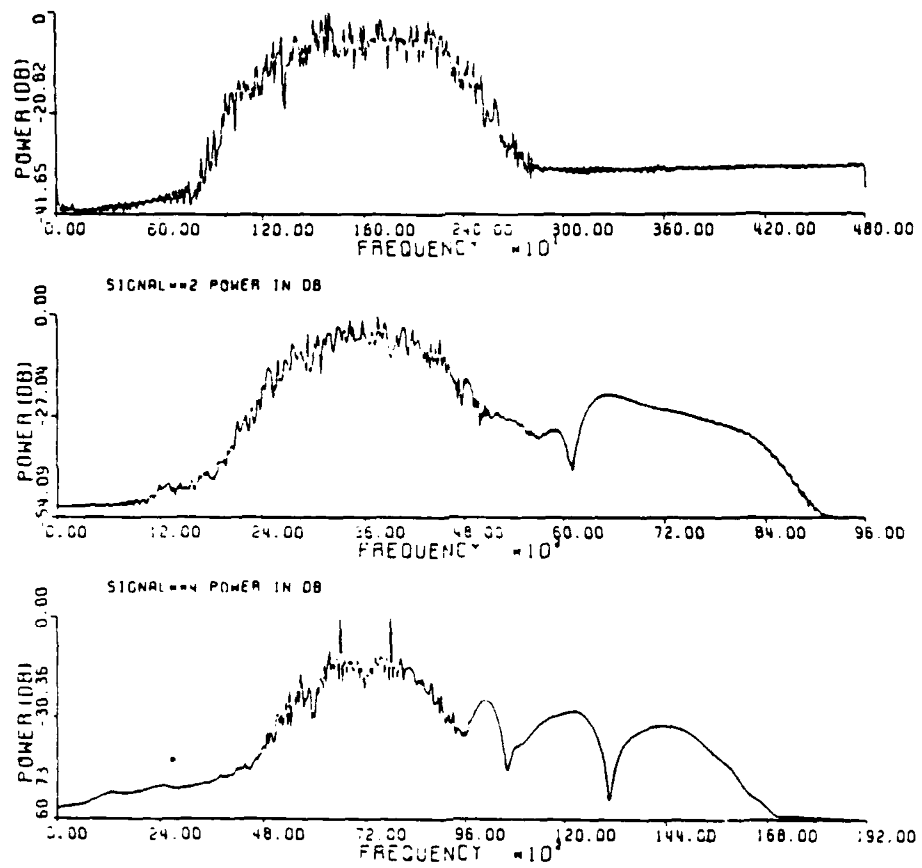


Figure 16. Spectra of powers of a CCITT V.26b signal.

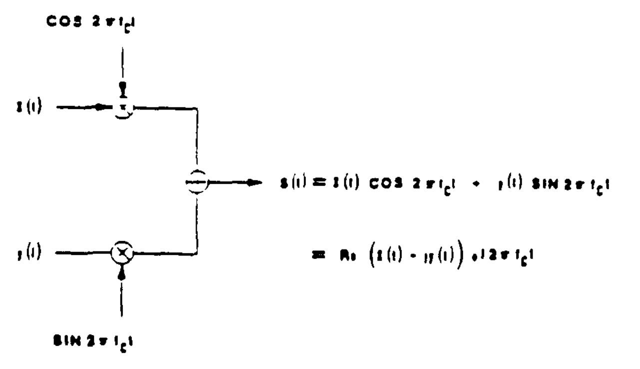


Figure 17. Model for generation of two-dimensional signals.

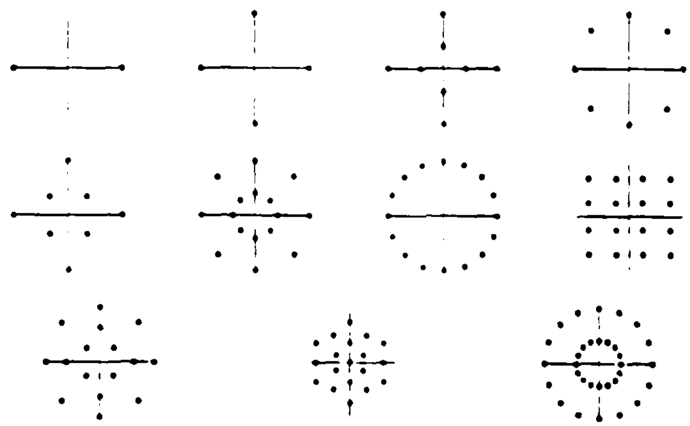


Figure 18. Various PSK and QAM constellations. From top, left-to-right: BPSK, QPSK, QASK, 8-PSK, V.29b, 8-PSK (two amplitudes), 16-PSK, 16-QAM, V.29, Codex 19-state, 16-PSK (two amplitudes).

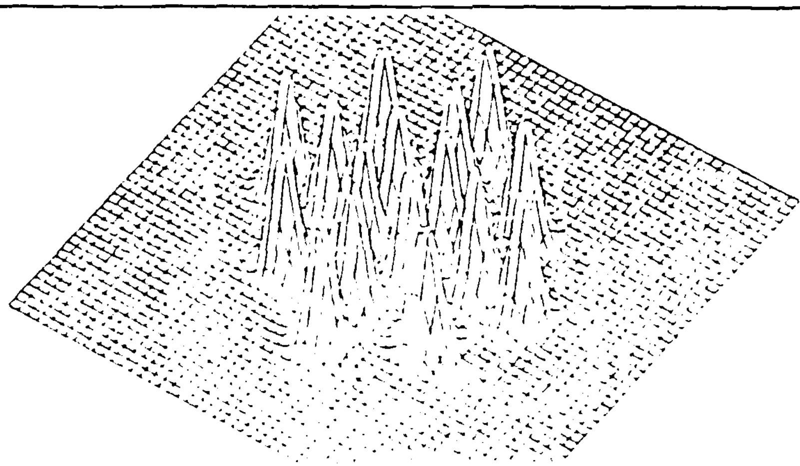


Figure 19a. 3-dimensional histogram of V.29 at 20 dB SNR.

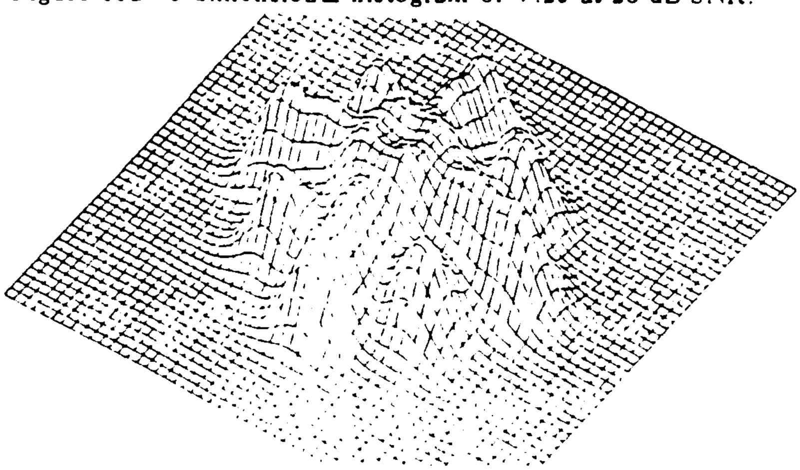


Figure 19b. 3-dimensional histogram of V.29 at 15 dB SNR.

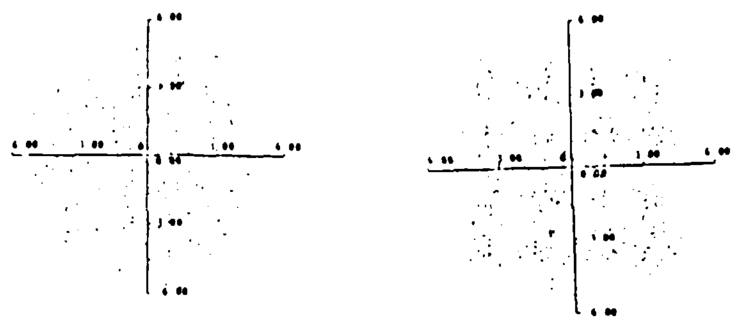


Figure 20. Amplitude-phase plots of a QAM signal before-and-after Godard blind equalization.

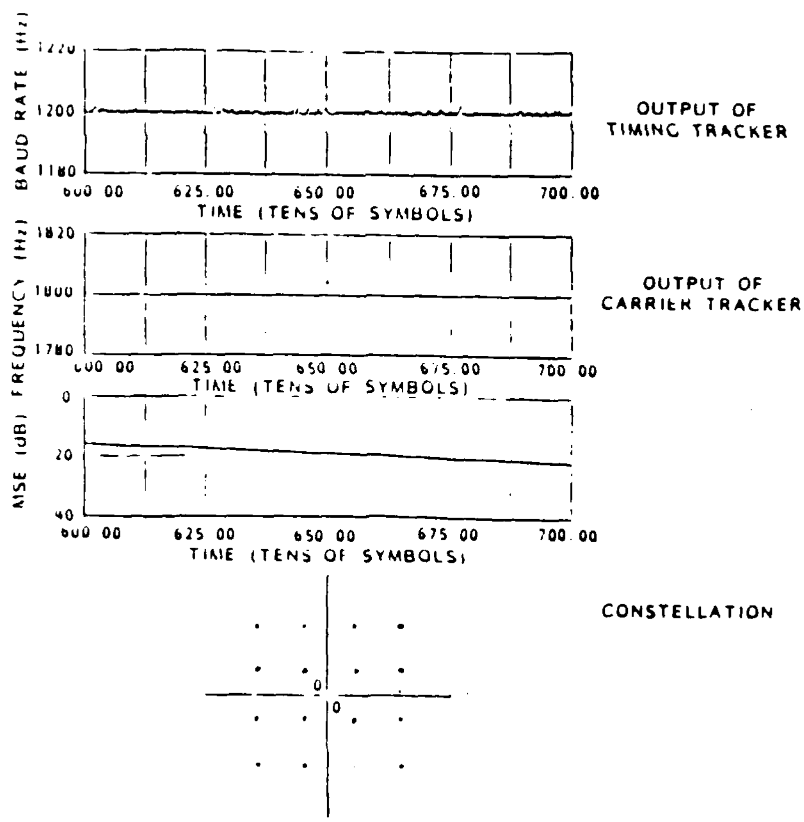


Figure 21. Demodulated 16-QAM using a decision-directed equalizer.

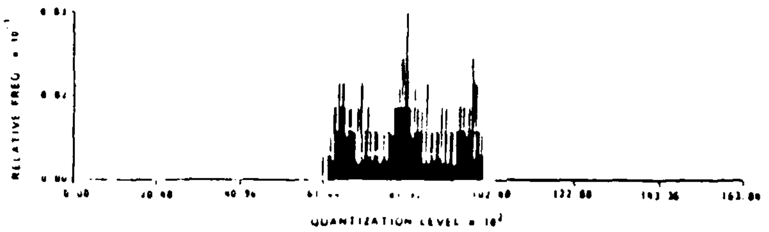


Figure 22. "Triangular" histogram of a basebanded duobinary signal.

PROPERTIES OF SIGNALS IN 4 kHz VGCs	VOICE	MCVFT	FSK	MSK	BPSK	PSK/VSB	QPSK	DQPSK	OQPSK	8-PSK	QAM	M-DUO/SSB	DUO/OPR	A-FAX
envelope fluctuations	yes													yes
spectral shape	spikey	peak-valley			2 peaks							log sin	log cos	
keying spike(s) in PSD of envelope/delay & mult		many	yes	yes	yes	weak (yes at baseband)	yes	yes	2xBW	yes	yes	NO	NO	
histogram	narrow		u-shaped									triangular at baseband		
spike(s) in PSD	yes		yes			Bell 203						often		
spike(s) in PSD of square				2	carrier	carrier								
spike(s) in PSD of 4th power				possibly			carrier	2	carrier			possibly 4xcarrier+BR		
spikes in PSD of 8th power										weak carrier	sometimes but weak			
instantaneous frequency			2-level			shows phase shifts								often constant

Figure 23. Summary of characteristics of various signal types.

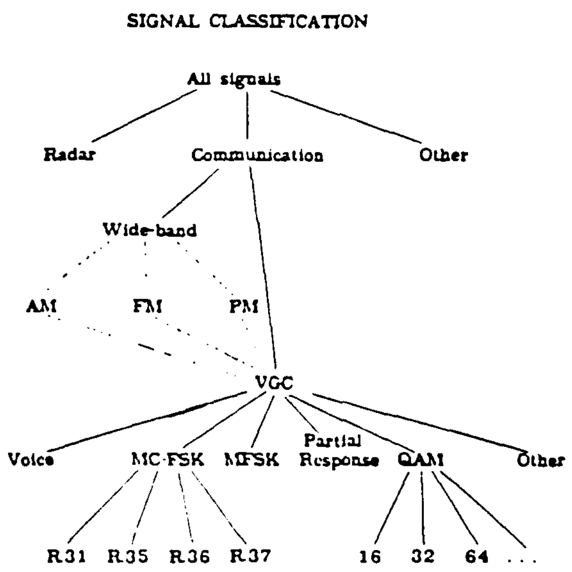


Figure 24. RBAS decision tree.

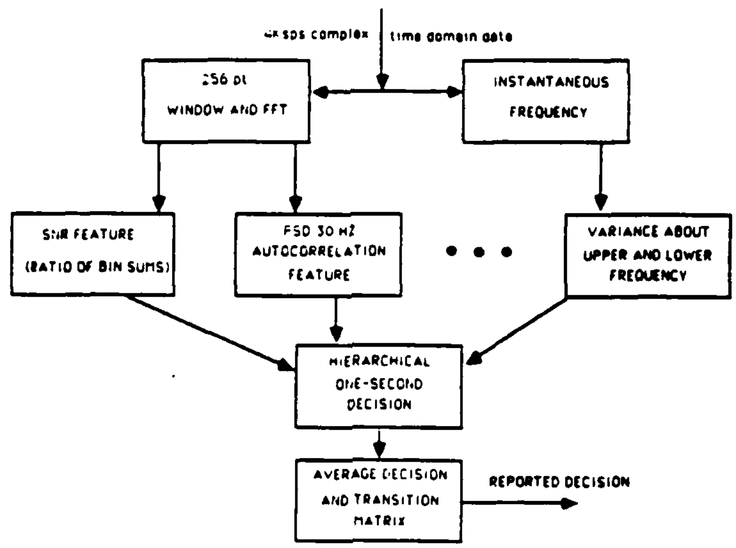


Figure 25. Block diagram of the "Fine Classifier."

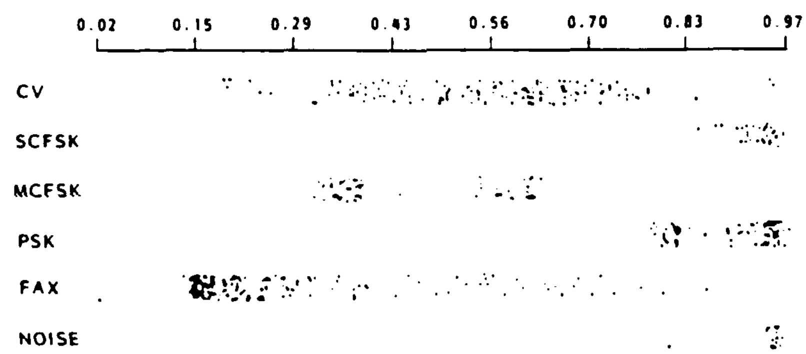


Figure 26. PSD Autocorrelation Lag Ratio.

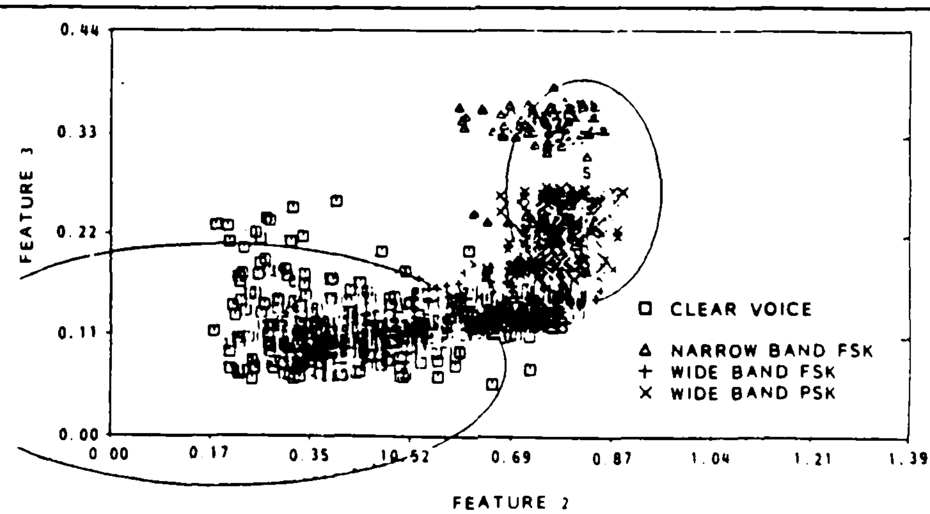


Figure 27. Two-dimensional scatter plot, showing the extent to which two features separate signal classes.

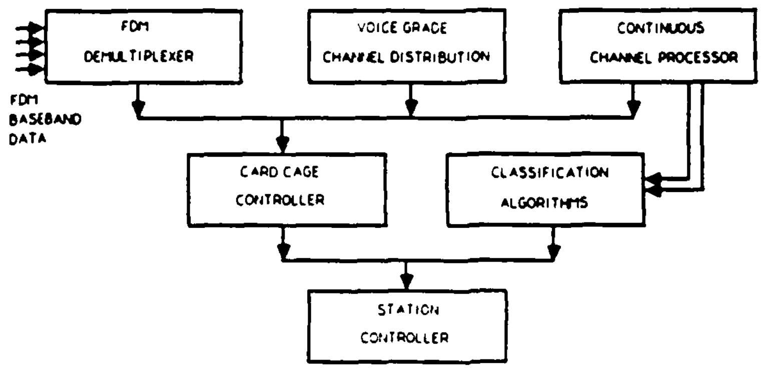


Figure 28. Lockheed's Voice Channel Processing Module.

Communication Channels

Session Chairman: William C. Lindsey

Phillip Bello
Radio Frequency Channels

Kenneth Wilson
Optical Channels

Paul Sass
Wideband Channel Measurement Experience

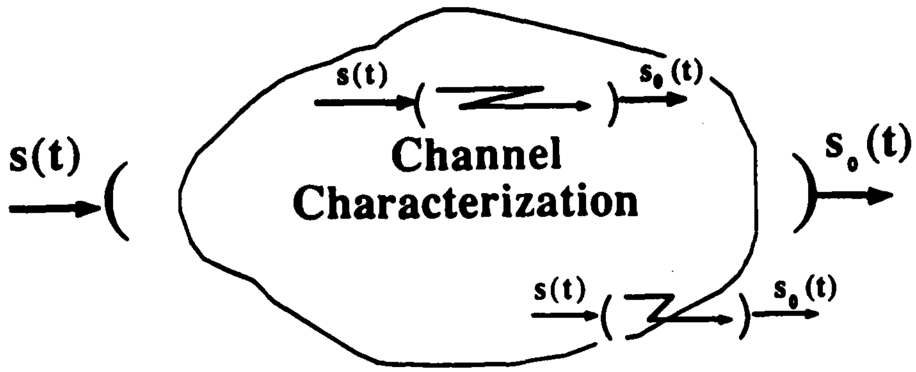
Allan Schneider
Delay Spread Estimation for Time-Invariant Random Media

ROBERT SCHOLTZ: I'd like to turn this afternoon session over now to Bill Lindsey. By the way, we will have a picture break during this session. I think it is good to get everybody together and take one group photo at these meetings. If you want a copy of it by any chance (we just save it for posterity), see Milly and she'll try to make sure that you get one. I think then we will try to do a coffee break on the fly because taking the picture will take a bit of time. Other than that if anyone wants to wander back and get something to eat or drink during the session, please feel free to do so.

WILLIAM LINDSEY: Thank you Bob, and hello again. It's nice to see so many attend our Workshop. It's nice to see old faces as well as new faces, too. Sometimes it is the ritual of the Session Chairman to start out with a joke. I have looked at all my viewgraphs and I couldn't find anything that was funny about them, so I don't have too much to say in this regard. I did look across the lake this morning at one of the Indian tepees and I see that they are still using the Huffman code to signal. Professors Lloyd Welch and Bob Peile are trying to get me to break early so that they can take a canoe across the way to see what they are doing in the way of adaptive coding!

Anyway, here's what the session is about today; I will start by giving you an overview. Perhaps an overview which addresses the problem of channel characterization from the perspective of what now I call the classical way. Then I'll apply this approach to a scintillation channel model in space and time. Secondly, Ken Wilson will also talk about an application to laser communications. We have Mr. Channel Characterizer himself, Dr. Phil Bello, who will pick up the second time slot on HF channels. After the break, we'll talk measurements. Then we will open the session up for discussion.

Bob Scholtz asked me to get involved in this session and I was delighted to do so because this problem area has been of longstanding interest to me, i.e., Channel Characterization. I have thought about it for a while and came up with a session theme which centers around what I'd call a class of *linear, random space-time varying channels.* Since we wanted to exclude notions related to modulation-coding questions, the question I ask is: Exclusive of any additive noise in the channel, what does the communication engineer need to know about the physical channel in order to design a combined modulation-coding scheme to mitigate the deleterious effects on communications performance? And so it is that I would like to address that issue and try to answer that question; whether or not it's *complete* in terms of solution is another issue. It is that, I believe, we do understand what we need and I'd like to talk about that from the perspective of a space-time channel model. Consider the point source at source point $(\hat{r}_s; t)$ at time t. I've illustrated this in CHART #3; an observed world point can be either in the near field or the far field, our interest in communications is generally in the far field. This sets up the geometry of the problem which leads me into a notion of a space-time channel. If I now take some current, some charge at that source point and move it around, I create an electric field at the observed point, this field will satisfy Maxwell's equations. Now in the absence of any turbulent media, I'll move the charge there and move it around, it turns out you create what we call the free-space field. It is that I would like to relate the free-space field (or the free-space electric field) to the observed field. It turns out that's easy enough



- Bill Lindsey : Overview: Space-Time Channels
  Scintillation Channels
- Phil Bello : HF Channels
- Ken Wilson : Optical Channels
- Break for group photograph
- Paul Sass : Wideband Channel Measurement Experience
- Allan Schneider : Delay Spread Estimation for Time-Invariant Random Media

CHART #1

SESSION THEME: CONSIDER THE CLASS OF **LINEAR, RANDOM, SPACE-TIME VARYING CHANNELS**

QUESTION: Exclusive of the additive noise, what does a Communications Engineer need to know about the channel in order to design combined modulation coding schemes to mitigate deleterious channel effects on bit error probability?

CHART #2

Illustrating the Geometry Between the Source and Observed World-Point

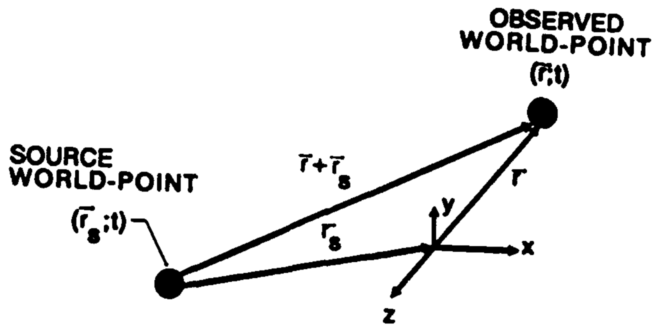


CHART #3

Space-Time Field Concepts and Geometry

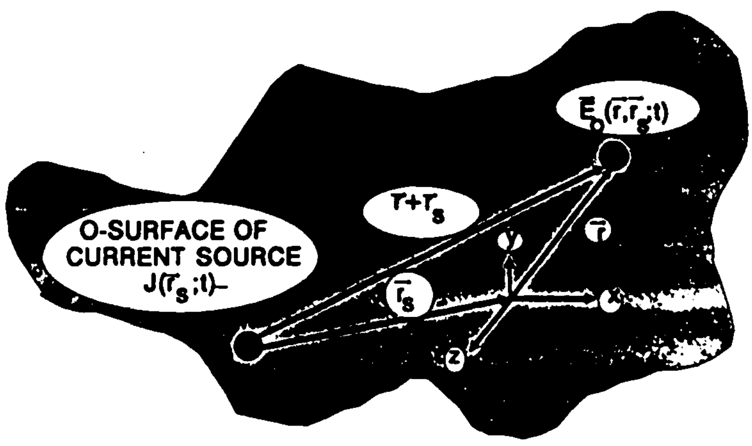


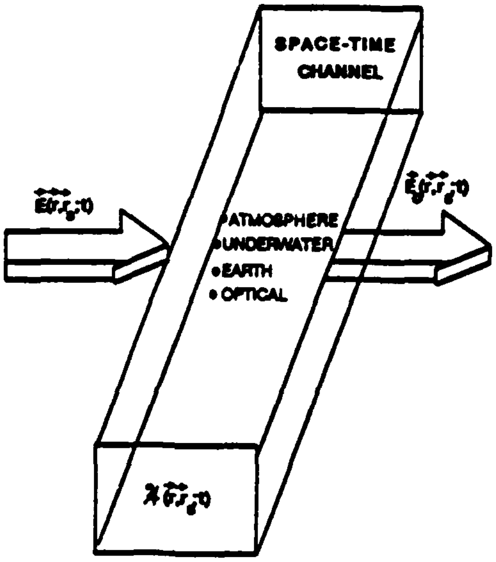
CHART #4

to do using Maxwell's equations. And when you play with them a little bit and write them in their integral form, you can actually show that everything we see in the real world is through a space-time filter. That's what this equation says at the top here (CHART #5), that is the free-space field is related to the output field in a turbulent media through a space-time filter. And we understand the space-time filter, it turns out that that it is characterized in terms of Green's diadic function and hence, as far as we're concerned, if we could get the free-space diadic Green's function and the Green's function in terms of a turbulent media through which the electromagnetic wave is traveling, then one has what I call now a *space-time channel transmittance function*, STCTF. The only "hook" then is how does one characterize this space-time filter? I am going to talk about this problem further.

Now if you take this equation and play with it in the far field, which is interesting to us with a point source, or you talk about a source continuum of points, then the STCTF still holds and the same problem as before is faced. The space-time channel transmittance function is much more complicated, however. But if you go into the far field then you have this kind [CHART #6] of an input-output representation which is interesting to communication engineers because they frequently look at the output of an IF strip and all they see is a voltage fluctuation versus time, and that's what this chart is saying. It shows how (in the far-field), the space-time input signal is related to the output signal; thus the output spectrum is a product of two spectral factors. There's one dealing with the spatial wave vector and time. Normally we deal in characterization of channels strictly with the temporal frequency variable resulting from the transformation of the time variable. We can see that it is related to the velocity of the medium and hence should be kept track of if it is, you want to try to tie the physics of what's going on in the medium, in terms of what I've called the spatial transform of the channel transmittance function. Now if you look at this for two cases, it's very interesting. Many times I have looked at Phil Bello's paper on channel characterization and I couldn't figure out what he did, but after a while I was able to see that, in fact, he showed that the channel is time-invariant, i.e., time-invariant in the sense that it doesn't depend upon time, and there's no such channel, but at least we'd like to think that there is. You can actually show that the output spectrum is related to the spectral response in wave number to the input spectrum through this equation. So it is a product of transforms. If on the other hand, you deal with the spatially-invariant channels, things that do not change with delay, then it turns out interestingly enough that the output signal is related to the product of the signals. So we see we have duality in channel response here whereas, when the input signal looks like a particle, the channel looks like a screen from the physics point of view with some of the particles falling through. On the other hand when it is time-invariant, the input signal looks like a wave and the output signal is related to the input signal through the convolutional integral. So in reality, what's happening is that something in between is always what is taking place. But it turns out that these two are interesting cases to study, and are of greatest practical interest.

Now the next (CHART #7) shows some relationships that will differ slightly from those that engineers typically use as performance characterization measures. It turns out if you

SYSTEM MODEL FOR SPACE-TIME CHANNEL



$$\overbrace{\vec{E}_0(\vec{r},\vec{r}_s;t)}^{\text{OUTPUT FIELD}} = \underbrace{\mathcal{H}^*(\vec{r},\vec{r}_s;t)}_{} \; \overbrace{\vec{E}(\vec{r};\vec{r}_s;t)}^{\text{INPUT FIELD}}$$

Space-Time Channel Transmittance Function = ST-CTF

$$\overbrace{\mathcal{H}^*(\vec{r},\vec{r}_s;t)} = \frac{\mu}{\nu_0} G_0^*(\vec{r},\vec{r}_s;t)G^{-1}(\vec{r},\vec{r}_s;t)$$

• $G(\vec{r},\vec{r}_s;t)$ - Green's Dyadic Functions

CHART #5

FAR-FIELD RESPONSE OF SPACE-TIME CHANNELS

$$s(t) \longrightarrow \boxed{h(\vec{r},\vec{r}_s,t)} \longrightarrow s_0(t)$$

$$s_0(t) = \int_{-\infty}^{\infty} S_0(\omega)\, e^{j\omega t} d\omega/2\pi$$

WHERE

$S_0(\omega) \triangleq H(j\vec{\kappa};t)S(\omega)$

TIME-INVARIANT → $S_0(\omega) = H(j\vec{\kappa})S(\omega)$

DUALITY IN CHANNEL RESPONSE

$\vec{\kappa} = (\omega/v_x, \omega/v_y, \omega/v_z)$

= VECTOR WAVE NUMBER

SPATIALLY-INVARIANT → $s_0(t) = h(t)s(t)$

$S_0(\omega) = H(j\omega) \circledast S(\omega)$

C 047

CHART #6

Space-Time Correlation Function

$$R_h(\Delta\vec{r},\vec{r}_s;t) \triangleq \langle h(\vec{r}_1,\vec{r}_s;t))h^*(\vec{r}_2,\vec{r}_s;t)\rangle$$

and

$$\rho_h(\Delta\vec{r},\vec{r}_s;\tau) = \frac{R_h(\Delta\vec{r},\vec{r}_s;t)}{R_h(\vec{0},\vec{0};0)}$$

ST CTF Correlation Parameters

NOTATION	DESCRIPTION	DEFINITION
τ_c	CORRELATION TIME	$\int_{-\infty}^{\infty} \rho_h(\Delta\vec{r}=\vec{0},\vec{r}_s;\tau)d\tau$
L_x	CORRELATION DISTANCE IN x	$\int_{-\infty}^{\infty} \rho_h(\Delta x,\Delta y=0,\Delta z=0,\vec{r}_s;0)d\Delta x$
L_y	CORRELATION DISTANCE IN y	$\int_{-\infty}^{\infty} \rho_h(\Delta x=0,\Delta y,\Delta z=0,\vec{r}_s;0)d\Delta y$
L_z	CORRELATION DISTANCE IN z	$\int_{-\infty}^{\infty} \rho_h(\Delta x=0,\Delta y=0,\Delta z,\vec{r}_s;0)d\Delta z$

CHART #7

JOINT SPACE-TIME AMPLITUDE AND PHASE PDF

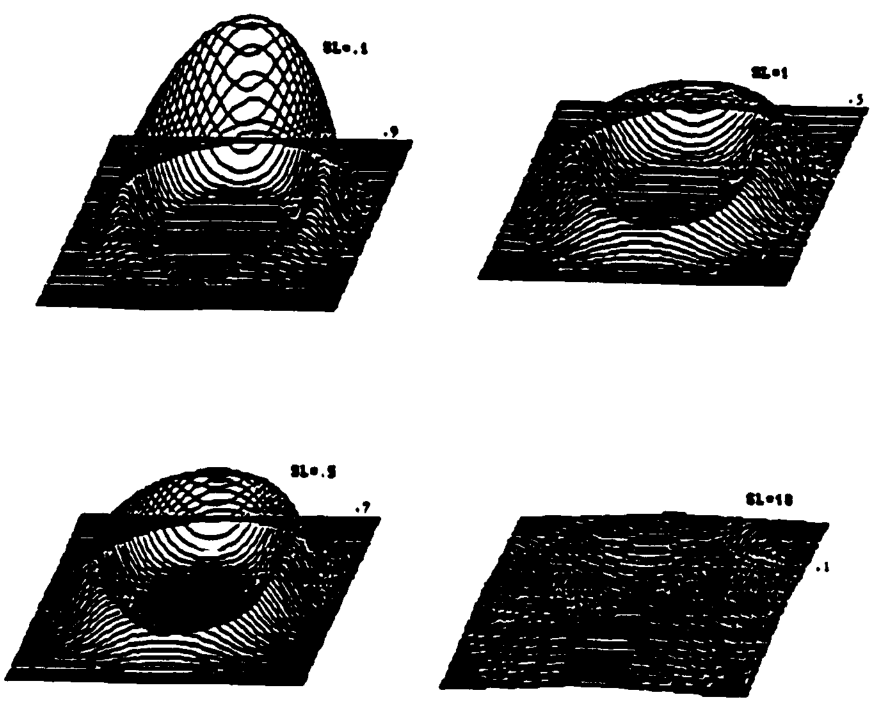


CHART #8

look at the space-time correlation function of the space-time channel transmittance function, you get a correlation function of this form and you can select and design some parameters that are related to the correlation time and correlation distance. Basically, the correlation time being the integral of the normalized correlation function, the same notion is indicated with the correlation distance. It's typically these things that you're familiar with.

I don't want to spend a great deal of time here. There are parameters which can be associated with the channel memory. Typically we talk about the coherence time as related to inversely what we call the Doppler spread of the channel. Then in the case of correlation distances, we've only dealt with just a delay spread in one direction and one has to normalize this by the velocity to get something related to the delay spread parameter which is inversely proportional to the coherence bandwidth. Now it is my desire for some time to relate these four parameters to what the radio physicists have been using for a number of years. By studying and playing with this for a long time, it turns out that there are some things that are interesting to do and I'd like to show you what they are with respect to an application to the scintillation channel. But in trying to answer the question posed earlier, it is my opinion that we need a characterization of the space-time channel which depends on the frequency band in the electromagnetic spectrum we are using. And it also means that we need a statistical characterization of the STCTF function, both in terms in the physics of the medium, whether it is friendly-hostile or combination environments, in which the geometry of the transmitter-receiver enters; it is the nature of this function which the electromagnetic wave is actually propagating.

Now I would like to apply this notion to the scintillation channel, then we'll take up the other channels. Now since the geometry is important, here is a particular type of geometry that I've looked at, viz., one with a satellite transmitting signal energy to an Earth station through a plasma that is located in the signal path. This will be what I call scenario number one. A second scenario is where the media is disturbed (it didn't come out too shaded), but the plasma is more diffuse, in some it's more layered, so that with a combination of layers the theory could be put together from the boundary conditions (at the various layers) can be met so that the amplitude statistics, phase statistics as well as the space-time spectrum here, can be obtained at the terminal. Now that you have the idea of the geometry, it is that I have placed the forward axis of the transmitting antenna along the z-axis. This simplifies what I am going to consider in the following discussion, viz., the case where energy is propagating in the z-direction. I'm not going to take into account the transference plane because it's too complicated to try to tell you what's going on in that problem. I will try to illustrate what it takes to characterize this channel from a physics point of view and from a communication engineering point of view. Here is an interesting CHART that I made up. It shows that various things drive the ionization process in the Earth's atmosphere. They include cosmic rays, solar rays, and in rare cases maybe there's a nuclear detonation that would disturb the media; this would also be of interest but I hope with zero probability. So I'm going to try to walk you through four or five steps that I've been through over a period of time. It shows one how to interconnect the physics of the particles within the media to

the channel transmittance function in space and time and then demonstrate for you some types of statistics that one can get out of this. The formulation goes like this. If you take the plasma theory and you take a look at it, it turns out that you can make a stochastic differential equation out of that theory and derive the Lorentz force equation. As a consequence of deriving the Lorentz force equation, you can come up with a description of a space-time ionization process driven by these sources. And from that you can write a permittivity tensor which is random in nature. And this permittivity tensor is what is needed to interconnect the physics of the medium to Maxwell's equations. When you play that game, after much work, it turns out you can write down some fairly interesting things in terms of the space-time electron density process. Once you have this permittivity tensor, it turns out you can formulate the space-time channel transmittance function in the following way. Now there are a lot of theories that people have looked at in the past. Those that deal with what I call weak fading or weak scintillation effects, this can be absorbed as a special case of a technique which was introduced by Stokes himself in 1890. I do not have time to tell you what happened between 1890 and 1950. In the 50's Professor Richard Bellman at USC took advantage of previous work and he formulated what he called an "invariant bedding technique" which you can apply to characterize the channel in terms of the physics of the medium.

Now, I don't have time to tell you much more about this other than give you some bottom lines as to what happens. It turns out that there are some interesting physical parameters that come out of the plasma theory and only a few of them that I'll relate to you because there are numerous others, which I've elected for reasons of good physical arguments, to ignore here. The electron density process, its standard deviation, the ambient electron density which is present in static situation, a scintillation strength which is a parameter for the statistics of the amplitude and phase-time process, the space-time normalized correlation function which is dependent upon the electron density, and the correlation distances. So you see, what drives these amplitude and phase processes are related to each other through stochastic differential equations. Moreover what makes them stochastic is the fact that you're driving the equations with the space-time electron density process. By normalizing things the parameter of great interest turns out to be this one which is characterized in terms of the variations in the density of the electron density. It is characterized as the function of the frequency in the electromagnetic spectrum that you're using, the so-called plasma frequency which is known as a function of such parameters as the charge on the electron, its mass and so on. In fact, I show it here. But here is the parameter of great interest. You can see that if the frequency at which you're operating is greatly above the plasma frequency that the effect of variations in the electron density in the path of the signal become inversely proportional to the frequency square. So that at lower frequencies, such at HF, it would seem to be a higher degree of fluctuations whereas if you go perhaps to the optical channel then this scintillation may not be as strong there. In fact, that is indeed true in practice.

I am now going to show you a comparison of the statistics of the amplitude and phase process for propagation scenario No. 1 as derived based on using the invariant bedding technique. And as a by-product of this you

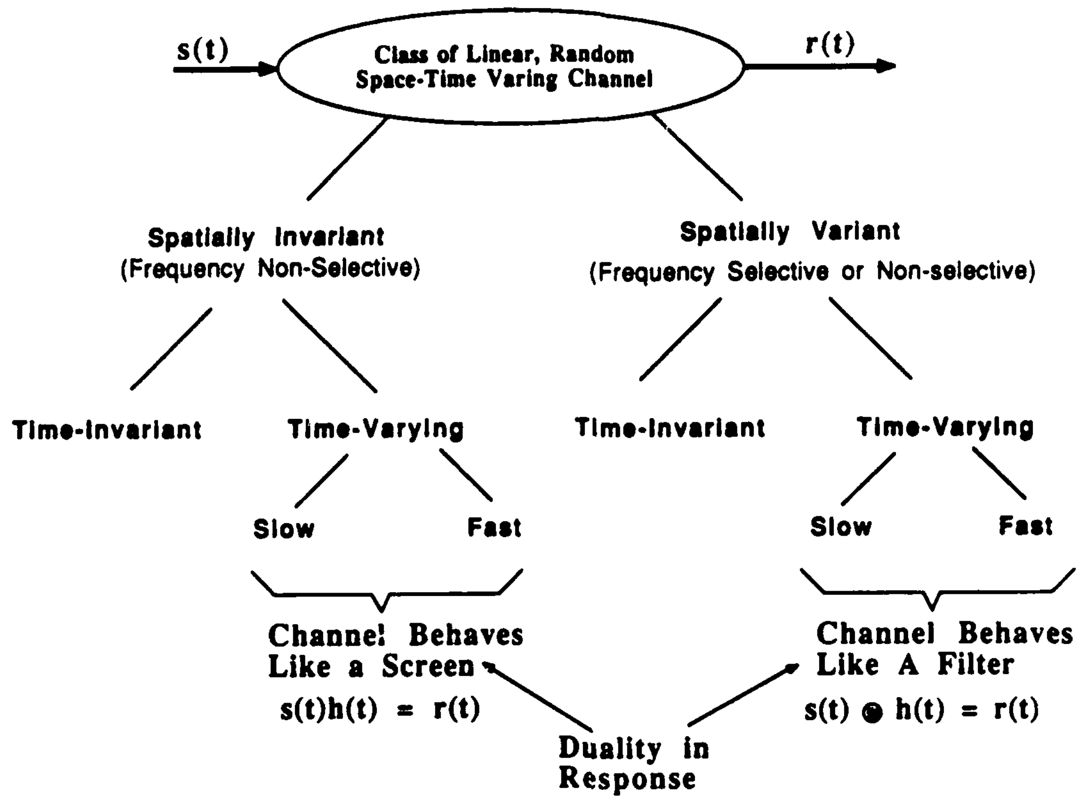


CHART #9

History

- Stokes - Huygens (circa 1890)
- Ambarzumian, Atmospheric Scattering, 1943
- Chandrasekhar; Principle of Invariance, 1950
- Bremmer; WKB Approximation *, 1951
- Bellman, Harris, Kalaba of USC; 1958

* WKB - Wentzel, Kramers, Brillouin

CHART #10

willsee other known solutions come out as special cases of a more general result. This is a very busy chart, but you can see that I've tried to tell a story. If I can walk you through this chart, starting at the top and then going to the bottom. The conditions of scintillation are going from strong to weak to none. The methodology used to arrive at the characterization of space-time amplitude process of the electromagnetic wave is the WKB method. Next we say the Rytov result, and then the result coming from the geometrical optics method. These are methods of what's called smooth perturbations, if it is you read books on radio-physics. It turns out that you play the game in solving this differential equation that I talked about, matching boundary conditions, then the amplitude process as a function of positions Z_1, Z_2 at time T looks like what is indicated here. The phase process, to a first-order, is always Gaussian and depends on the integrated electron density along the propagation path at time T. This is known. So the phase statistics are always Gaussian. Now if we look down this path indicated here, we go to sort of strong scintillation. You get this kind of a function. This is driven, as you recall, by the electron density process. If you look at the weaker scintillation case, this becomes root Cauchy, where these are kind of the ratio of two Gaussian processes. And then on even weak scintillation, Rytov's method, which leads to log-normal statistics, is known and hence this function reduces to that. And finally the geometrical optics method accounts for the fluctuations of the phase process on the electromagnetic wave. Of course, in the limit as you turn the "scintillation off," then the amplitude statistics become deterministic with no phase variation at the output. If scintillation is strong then the amplitude density goes to the origin, and the phase is uniformly distributed.

Now I will show you some pictures and show you what the first order of phase-time statistics look like for this process. Now you have to think in three dimensions even though this is in a plane. Here is a statistical characterization of the amplitude process associated with the radio wave in the media. I will try to explain: This is probability axis and this is the scintillation strength (or this is time, if you will, after an event disturbs the ionosphere). For each one of these curves there are two parameters associated with it; they relate the physics of the media to the parameters that I'm going to talk about. Notice that if there's no scintillation then this is a delta-function. Now the two parameters that are here are the space-time correlation and the scintillation strength. If the scintillation is weak, you would expect the amplitude and phase processes to be highly correlated at the input and output of the two boundaries which define the turbulence region. For 0.9 correlation we see that the amplitude becomes something like this. Now then, as I turn up the intensity of the electron density fluctuations, the spatial decorrelation reduces with time and the scintillation strength comes up. And as I continue to increase the intensity of the electron density variations, the decorrelation or the correlation in time decreases, and of course this goes up, all to such a point that you can see that what is happening is that the mean value of the signal is moving to zero. It looks like the variance is about the same with regard to this particular density. And, of course, in the limit the amplitude density becomes a delta-function which lines at the origin. Now this density is unlike anything that I've worked; usually one assumes a Rayleigh or Rician statistic. It turns out that when

you turn the channel variations in a Rayleigh density off, it becomes a delta-function at the origin. It says that the channel is not random even though there's no amplitude out of the channel. So you must introduce a specular term into the model to avoid this. But these are the distributions in space and time. And then I have some three-dimensional plots to show you the combined amplitude-phase process. This starts out with $t = 0$, little scintillation, highly correlated input-output signal. As one turns up the scintillation, the space-time correlation is decorrelated and the probability mass diffuses and spreads until it is not an interesting channel to communicate over.

Finally, for purposes of completing my discussion, I'd like to compare (under the same conditions in terms of variance) the so-called Rayleigh distribution with a comparable one which comes out of the invariant imbedding solution. OK, that's all I have to say with regard to my discussion. I would like to turn the discussion over to Phil Bello who will talk to us regarding the HF channel.

PHIL BELLO: *Radio Frequency Channels*

This title seems kind of funny. It says Wideband HF Channel Modeling for a particular modem design. I mean, we should have a model of the channel which has nothing to do with the particular modem. But I just put it that way because that's how I got into HF channel modeling. At MITRE we have a modem and we are developing various aspects of it, and in order to come up with a design or optimize it, I had to have some models. So I'll just take you through my own thought processes as to how I came up with some models, and then you can say at the end that now you have been brought up to my own level of ignorance.

Here is an outline of my talk. [VIEWGRAPH #2] I'm going to give what I call a system engineer's or a black-box view of the wideband HF channel. And I'll show you what a link might look like utilizing a direct sequence spread spectrum modem. I'll get into propagation channel models and say something about additive disturbances. I had to cut down the number of my slides because I misunderstood how much time I had. I'll be flipping rapidly by some of them. Don't let that bother you because there's always a subliminal effect. [LAUGHTER]

[VIEWGRAPH #3] Let me say that MITRE has been involved in four research areas connected with wideband HF modem design. By wideband we mean of the order of 1MHz. Typically, as you may know, the bandwidths used at HF are 3KHz or less. We're involved in experimental modem design, measuring the impulse response of this time-varying HF channel, modeling the noise, and simulating in real time the wideband HF channel (MITRE Washington).

Here's sort of a system engineer's view of the bandlimited HF skywave channel. I want to spend a few minutes on this particular viewgraph. [VIEWGRAPH #4] I assume you are all aware that the HF channel is a multimodel channel and that there are reflections off the ionospheric layer and so on. So each one of these reflections or modes can be modeled by A random time-varying linear channel. The signal carrier frequencies of HF range from 3 to 30 MHz. The noise is rather severe for a number of reasons. We have narrowband communications interference. All those communicators are spread over the HF band and while you're trying to communicate with a 1MHz bandwidth, you get a large number of them sitting on top of you. Then we have man-made noise which can be severe depending on where you are located. Also there

WIDEBAND HF CHANNEL MODELING FOR DIRECT SEQUENCE SPREAD SPECTRUM MODEM DESIGN

Phillip A. Bello
The MITRE Corporation

presented at

ADVANCED COMMUNICATION PROCESSING TECHNIQUES WORKSHOP

14-18 May 1989

MITRE

VIEWGRAPH #1

OUTLINE

- A SYSTEM ENGINEER'S VIEW OF THE WIDEBAND HF CHANNEL
- A WIDEBAND HF DIRECT SEQUENCE SPREAD SPECTRUM LINK
- PROPAGATION CHANNEL MODELS
- ADDITIVE DISTURBANCES

VIEWGRAPH #2

MITRE

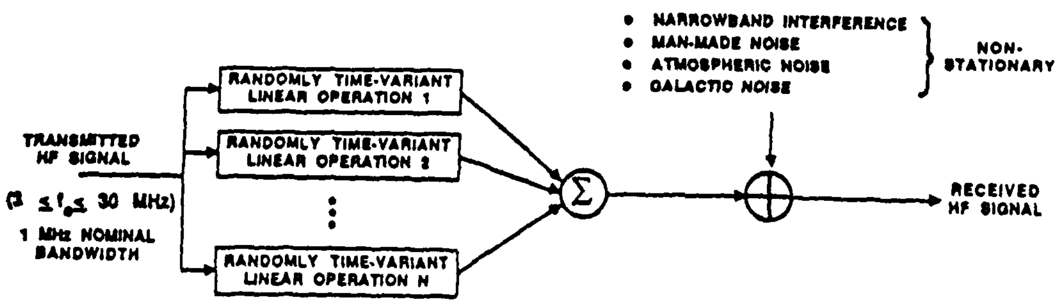
MITRE R&D EFFORTS IN WIDEBAND HF SPREAD SPECTRUM COMMUNICATIONS

- EXPERIMENTAL MODEM DESIGN
- BANDLIMITED IMPULSE RESPONSE MEASUREMENTS
- NOISE MEASUREMENTS AND MODELING
- REAL TIME CHANNEL SIMULATION

VIEWGRAPH #3

MITRE

A SYSTEM ENGINEER'S VIEW OF THE BANDLIMITED HF SKYWAVE CHANNEL



APPROXIMATE TWO-STATE CHANNEL

- NORMAL OR UNDISTURBED (SMOOTH IONOSPHERE)
- DISTURBED (IRREGULAR IONOSPHERE, $\overline{(\Delta N/N)^2}$ SUFFICIENTLY LARGE)

NON-STATIONARY

VIEWGRAPH #4

MITRE

is atmospheric noise caused by lightning and intergalactic noise. Receiver noise isn't even up there, it's so small compared to everything else. But the important thing to note is that the total additive disturbances are non-stationary, unlike the additive white Gaussian noise channel. Unfortunately, the HF channel has this property of being very non-stationary, and in modem design and signal detection processing you have to worry about that. As far as the propagation medium itself, you can approximate that roughly as a two-state channel. The normal or undisturbed state is a smooth ionosphere. That is to say, the electron density, when plotted as a function of spatial coordinates has a smooth variation. The other state called disturbed, is due to an irregular ionosphere. The mean squared value of the percentage fluctuation in electron density $\overline{(\Delta N/N)^2}$, is a critical parameter in determining the state of the channel. This parameter relates to one that Bill Lindsey mentioned. Depending on propagation conditions $(\Delta N/N)^2$ can vary 10 orders of magnitude. And the transition between the normal state and the disturbed state is only 1 order of magnitude. That is why I say that it's almost like a two-state channel. Of course as the fluctuations in the electron density get worse, the channel gets more disturbed, but it is still classed as a disturbed channel. And what I'd like to do is present to you models of the non-disturbed channel and of the disturbed channels, and some measurements which support these models. I can't go through all the slides unfortunately.

Again, the propagation medium is non-stationary also. So here we have non-stationary channel, non-stationary noise, and what the devil do we do about it. How do we design a modem under these conditions? It's a puzzle. Well the potential solution to this problem is, we introduce a term called quasi-stationarity. [VIEWGRAPH #5] We hope that we can define the propagation medium by a set of parameters and then if we freeze those parameters we can define a set of statistics for input-output behavior which are well defined and with which we can do all kinds of wonderful analysis. We could define how a modem operates, and how to optimize its design. But then those parameters contain the non-stationarity, and we let those parameters slowly vary. The essence of our solution is to postulate quasi-stationarity and to hope that it's really true. It seems to be reasonably good, but time will tell. We can get our parameters for this channel either of two ways. We can do channel sounding or we can try propagational theoretic analyses. Bill Lindsey was introducing the ultimate in propagational theoretic analysis. He has a solution to all radio communications problems in his formulation. So this is a small corner of that. It's actually a fascinating theory as Bill is finding out. There has been some work done though, and I'm going to show you some theoretical predictions for the disturbed channel and compare them with actual measurements if I can get to them within my time allotment.

So what you do is postulate a stationary model with slowly varying parameters and then you do your design, build you modem, and then you collect statistics on these parameters. But to get the statistics you've got to measure all day, different seasons, throughout the year. In fact, it's a life-time job but you've got to keep doing it.

Here's [VIEWGRAPH #6] a potential block diagram for a spread spectrum direct sequence modem used over the HF channel. It's got coding and interleaving, frequency conversion, the usual things. The important thing I want you to look at is called a narrow-

QUASI-STATIONARY RADIO PROPAGATION CHANNEL MODELING

- CHANNEL CONTAINS ONE OR MORE "BLACK BOXES" DESCRIBED BY RANDOMLY VARYING SYSTEM FUNCTIONS
- MODEL EACH SYSTEM FUNCTION WITH MINIMUM NUMBER OF PARAMETERS BY

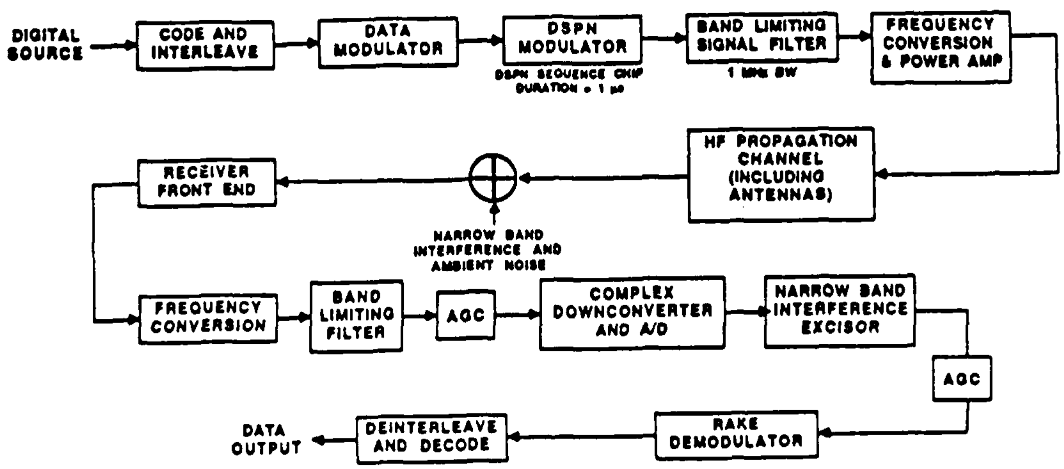
  CHANNEL SOUNDING

  PROPAGATION-THEORETIC ANALYSES
- POSTULATE STATIONARY MODEL WITH SLOWLY VARYING PARAMETERS
- COLLECT STATISTICS ON PARAMETERS

VIEWGRAPH #5

MITRE

SIMPLIFIED BLOCK DIAGRAM OF LINK PROCESSING



VIEWGRAPH #6

MITRE

band interference excisor. The reason that that's necessary is, as I pointed out before, that the HF channel's got this tremendous number of interferers sitting on top of you. You can remove them by some clever processing. It's an essential ingredient if you're going to have any hope at all of doing wideband HF communications.

I've shown a box called a rake demodulator. I don't know if you're familiar with the Price and Green implementation of adaptive matched filter-receiver for multipath channel. It's called a rake demodulator. This is the kind of thing that I've analyzed for the particular kinds of quasi-stationary models that I'm going to show you. There is not time to go into the structure of that particular modem. [VIEWGRAPH #7] That was one of those subliminal ones that will stand you in good stead later.

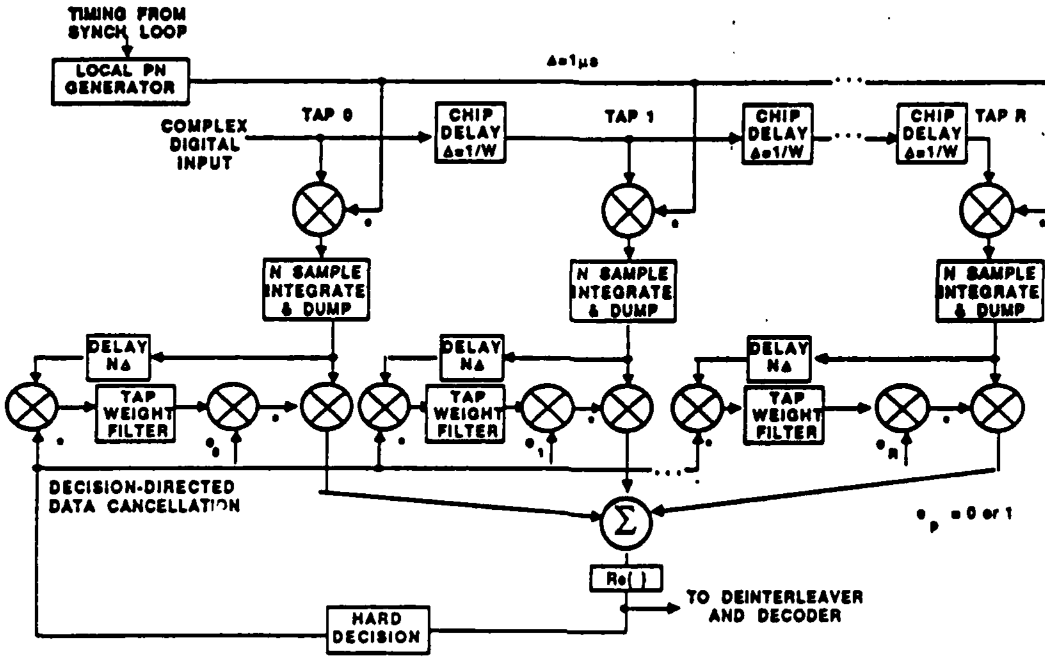
VIEWGRAPH #8 shows multiple reflections off the ionosphere. VIEWGRAPH #9 shows you how the various layer heights change with time. And the reason I'm showing that is I want you to understand that during the day, as the layers go up and down, that vertical motion for the undisturbed channel produces a Doppler shift. Mode Doppler shift happens to be an important parameter in determining the performance of the rake modem which tries to combine energy coherently and adaptively. To do this it needs to measure the channel impulse response. Doppler shift is a parameter that we have to collect more data on. I'll show you some measurements later.

This is an ionogram. [VIEWGRAPH #10] You may all be familiar with it, I don't know. Just give it a couple of minutes. What you do is transmit a pulse and vary the frequency of the pulse and see what you get coming back. And what you see is the energy coming back one or more delays with the delays changing with frequency. The different traces correspond to 1 hop, 2 hops, 3 hops path and so on. These may be further decomposed into high-ray and low-ray, but I can't get into that here. The important thing is that this is the ionogram for a *non-disturbed* channel. It is a relatively clean trace. Now I'm going to show you what the ionogram of the disturbed channel looks like.

You see what happens (see VIEWGRAPH #11) is you get spreading this way (vertically in delay). You had a thin line before but now you've got spreading in the delay direction. That's the property of the disturbed channel.

VIEWGRAPH #12 shows the multi-mode representation of the channel, the old standby in terms of modeling, time-varying disturbed channels and tap-delay lines. What you can say is that any linear, time-varying channel can be represented parametrically as a tapped-delay line in which the taps are spaced 1 over the bandwidth apart [VIEWGRAPH #13]. So you could say, well, great, now you've got your model, you have parameterized it. Unfortunately that won't get you too far, because the multipath spreads can be hundreds of microseconds and those taps are 1 microsecond apart since you have a bandwidth of 1 MHZ. So those time-varying complex coefficients may be hundreds in number. What are you going to do with those hundreds of coefficients? How are you going to characterize them? It's not too satisfactory. So you have to go and dig a little bit deeper to model this channel. Well, you start by looking at an idealized version of the ionogram [VIEWGRAPH #14] for 1 propagation mode. You look at that and what it shows is that the group delay as a function of frequency is approximately linear over 1 MHz bandwidth, but there's some non-linearity in

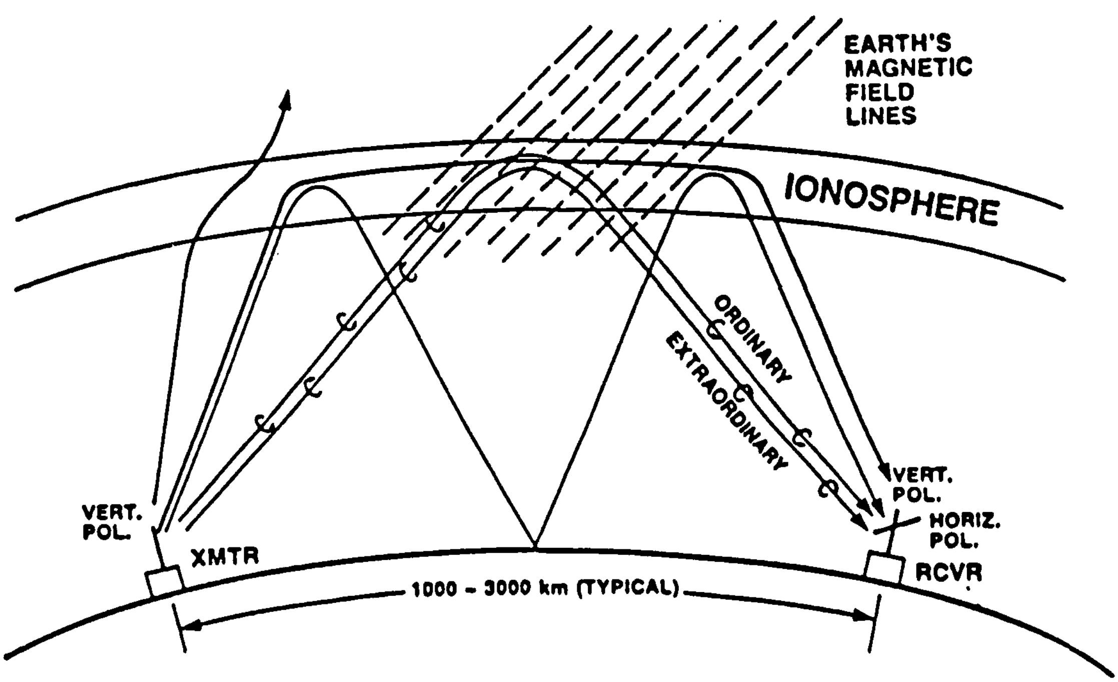
A DECISION-DIRECTED COHERENT RAKE PROCESSOR



MITRE

VIEWGRAPH #7

SKYWAVE HF CHANNEL

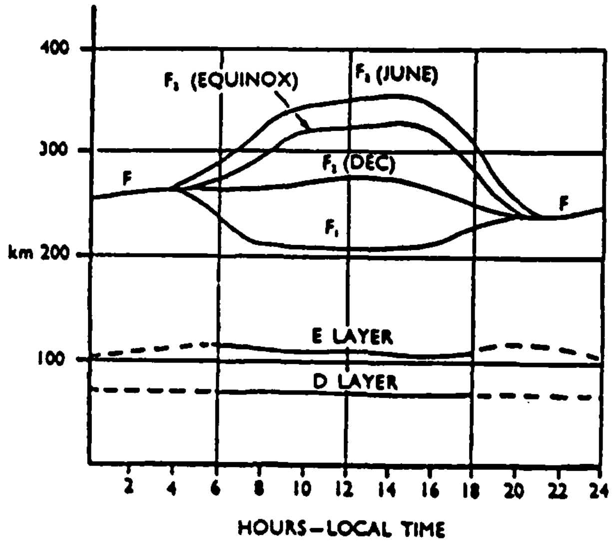


MITRE

VIEWGRAPH #8

TYPICAL DAILY VARIATIONS IN LAYER HEIGHTS FOR SUMMER, WINTER AND THE EQUINOXES IN MID-LATITUDES OF THE NORTHERN HEMISPHERE

F. R. East, "The Properties of Ionosphere Which Offset HF Transmission", from Point-to-Point Telecommunications, Febr. 1965, pp. 5-22



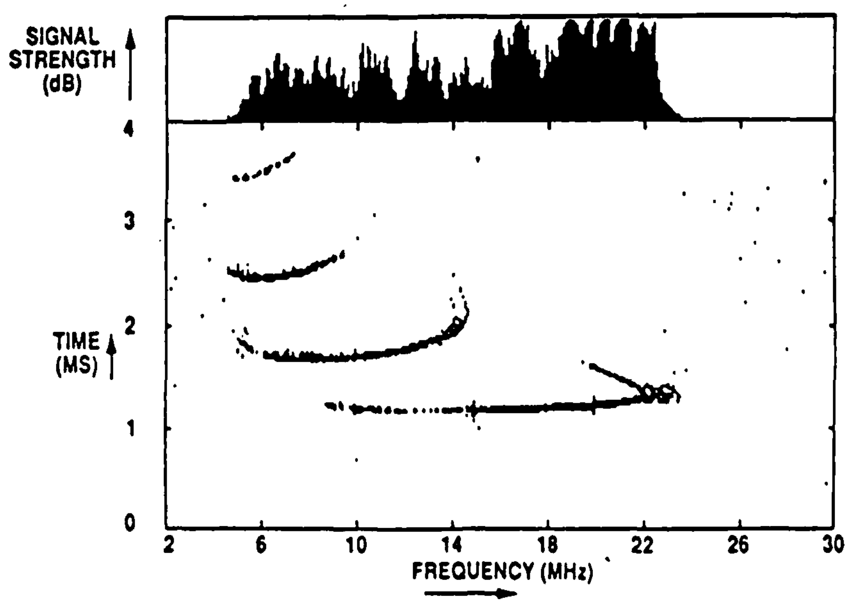
MITRE

VIEWGRAPH #9

IONOGRAM

Homestead AFB, FL to Bedford, MA

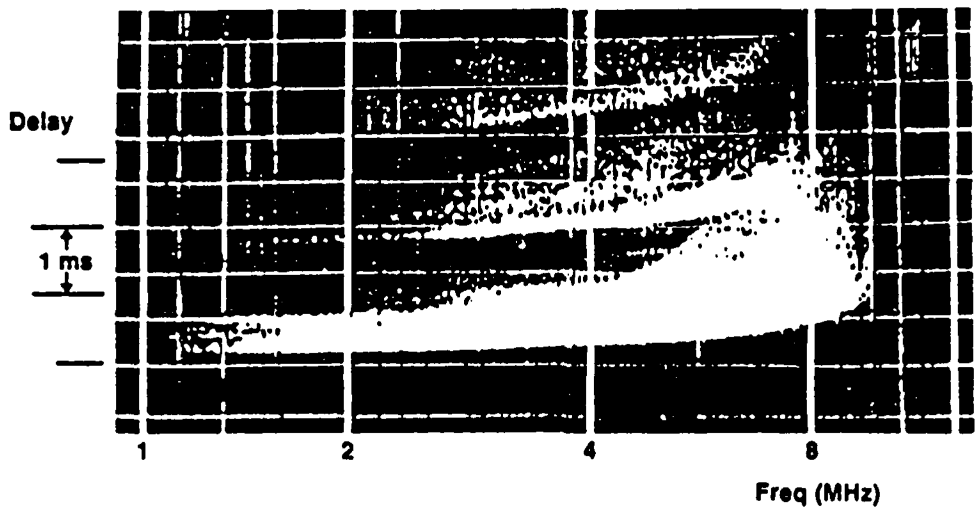
October 13, 1988 from 1200Z



MITRE

VIEWGRAPH #10

SPREAD-F IONOGRAM

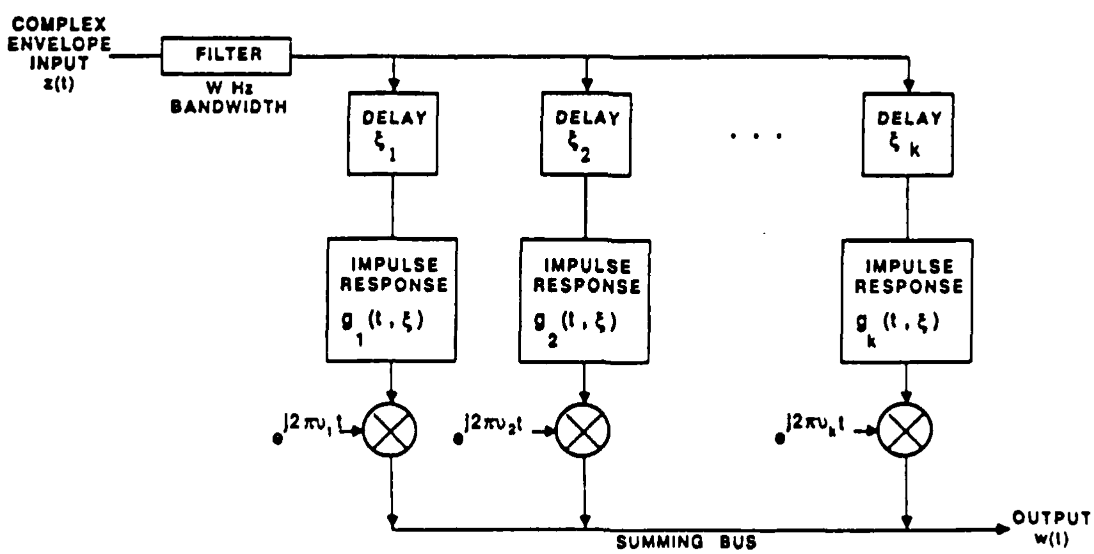


From King [Jour. Atmos. and Terr. Physics, 1970]

MITRE

VIEWGRAPH #11

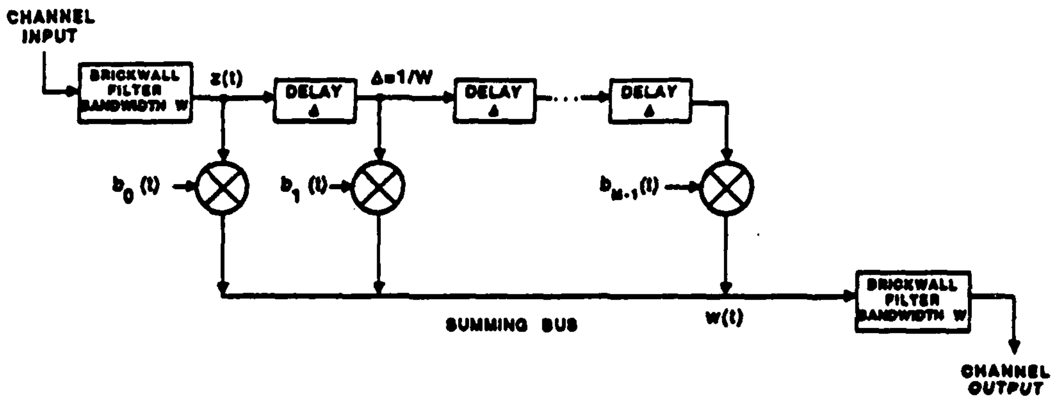
GENERAL MULTIMODAL REPRESENTATION OF WIDEBAND HF CHANNEL



MITRE

VIEWGRAPH #12

TAPPED DELAY LINE MODEL FOR BAND LIMITED WBHF CHANNEL



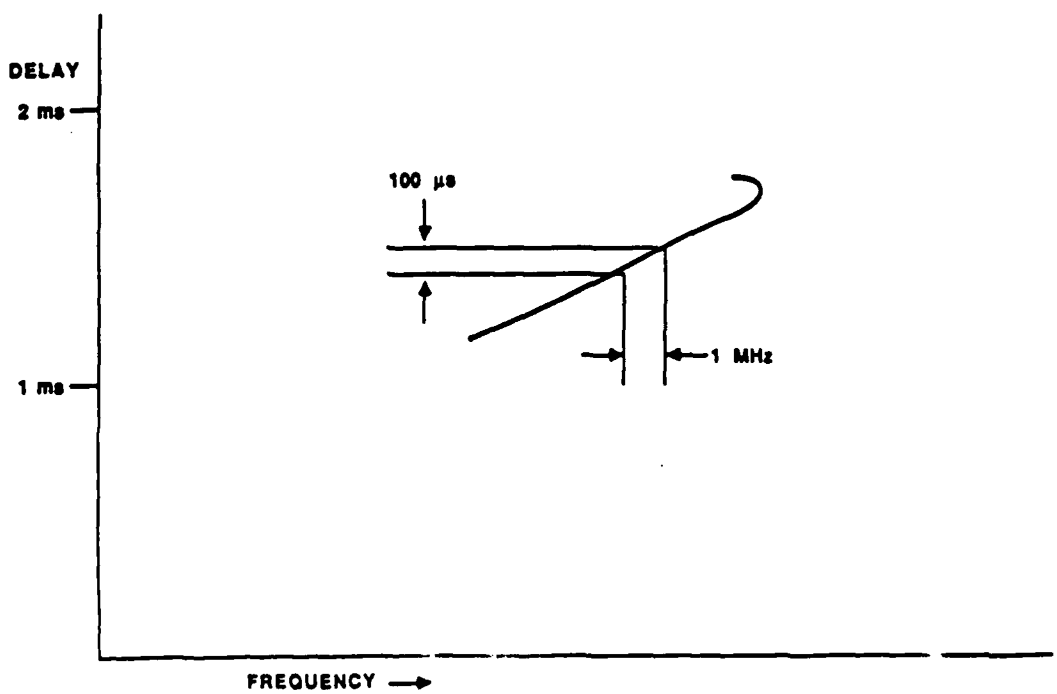
COMPLEX TAP WEIGHTS $b_K(t)$ DETERMINE CHANNEL CHARACTERISTICS.

W = BANDWIDTH OF TRANSMITTED SIGNAL

MITRE

VIEWGRAPH #13

IDEALIZED IONOGRAM EXAMPLE FOR NON-DISTURBED CHANNEL



MITRE

VIEWGRAPH #14

there. Your first concept is, OK. For the non-disturbed channel, I can model this as if it were a filter of 1 MHz bandwidth with a non-linear group delay distortion. The other thing to recall is that I told you that the layers are moving with time. What's happening is that that's moving, and because it's moving you've got a Doppler shift. So that what we have is a dispersive filter, a slow time-varying gain, and the slow time-varying Doppler shift. That's what the model is over here [VIEWGRAPH #15]. This is our quasi-stationary model for the non-disturbed channel: slowly changing delay, a dispersive filter, with hundreds of microseconds of delay possibly, dispersion, slowly-varying Doppler shift. Now I've made some calculations as to how a rake modem would work assuming that you had a white noise background (it could be non-stationary). And what you find is that the short-term performance, i.e., the quasi-stationary performance with the parameters fixed, depend primarily on the received power of all the taps in the tapped-delay line model. It depends on the delay spread that you need to accommodate in the rake combiner. It also depends on the Doppler shift. There is a Rome Air Development Center report (RADC-TR-98-91) which has all the theoretical background to it, but there's a 1988 MILCOM paper which has performance estimates which I'm not going to present here.

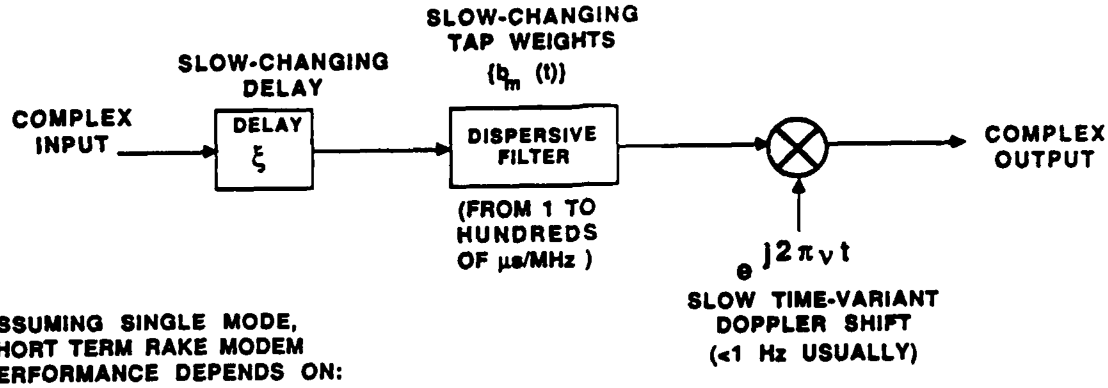
What you've got here are some measurements taken at MITRE recently. [VIEWGRAPH #18] We have a 1 MHz direct sequence modem and we can measure the energy in all those taps in the tapped-delay line channel model. Once again, we don't have time to discuss all those measurements. But you look at the time axis, you've got minutes on the bottom and the fluctuation rate of this total energy of this 1 MHz bandwidth for the propagation mode is slow. In fact, the correlation time is around 45 seconds, so we truly have a slowly-varying power in the 1 MHz bandwidth. Thus the idea of quasi-stationarity is not so far fetched. In other words, we do have a dispersive channel but the total power received over a 1 MHz bandwidth is varying slowly. Here's another example. [VIEWGRAPH #19] In this case the HF medium didn't cooperate. There was a solar flare and we had slow fluctuation, but all of a sudden we're right down to the noise floor. So we can't communicate at all under those conditions.

These [VIEWGRAPH #20] are probably distributions of the signal power in a 1 MHz bandwidth, comparing to the Rayleigh distribution. And you can see there's quite a bit of variation, but it's not as bad as the Rayleigh distribution. Here's some collected measurements [VIEWGRAPH #21] These were done at MITRE again. There will be another paper at MILCOM by Dan Perry discussing these measurements. There are 38 runs, 35 minutes each, covering some time span, but who can say that that defines what the HF channel is going to do. At any rate, you see the delay spread is on the right hand column, varied from 2 to 35 microseconds, the Doppler shifts are less than a Hertz, and so on.

What I'm now going to show you are some measurements taken years ago by SRI in which they are primarily concerned with frequency spread and Doppler shift. They had two paths [VIEWGRAPH #22]. Normally this path from Palo Alto to Fort Monmouth is a well-behaved, usually non-disturbed channel. The one from Palo Alto to Thule, Greenland is usually quite disturbed because it goes through the Aurora. They didn't have the hardware to measure the full impulse response so they just transmitted carriers and

QUASI-STATIONARY MODEL FOR NON-DISTURBED 1 MHz BANDWIDTH PROPAGATION MODE

(USUALLY APPLICABLE TO MID-LATITUDE PATHS)



ASSUMING SINGLE MODE, SHORT TERM RAKE MODEM PERFORMANCE DEPENDS ON:

RECEIVED POWER $\sim \sum_{m=0}^{M} |b_m(t)|^2$

DELAY SPREAD $\sim$ MΔ µs

DOPPLER SHIFT

PERFORMANCE PREDICTIONS:

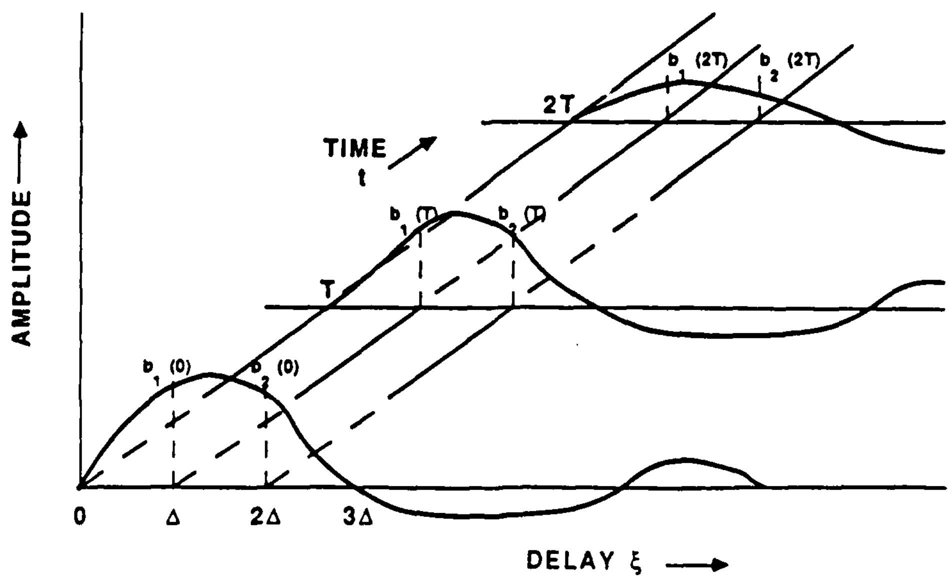
"Performance of Some Rake Modems Over the Non-Disturbed Wide Band HF Channel", P. Bello, MILCOM'88 Conference Record, IEEE/DoD/AFCEA.

"Performance of Four Rake Modems Over the Non-Disturbed Wideband HF Channel", P. Bello, RADC-TR-89-91, March 1989, (F19628-86-C-0001)

MITRE

VIEWGRAPH #15

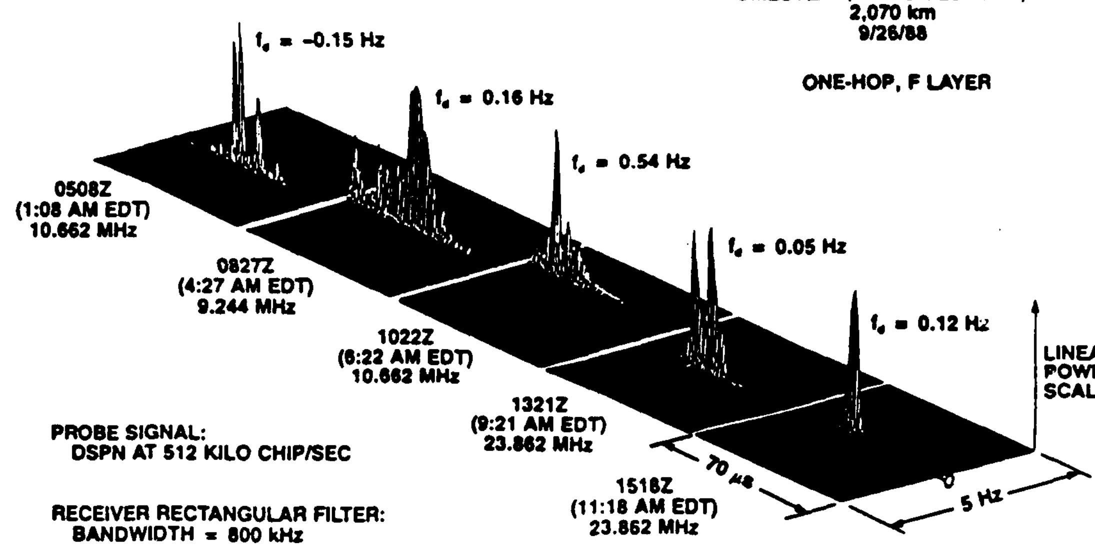
ILLUSTRATION OF MEASURED BANDLIMITED IMPULSE RESPONSE SNAPSHOTS



MITRE

VIEWGRAPH #16

WIDEBAND HF DOPPLER SHIFT AND TIME DELAY SPREAD DISPLAYS

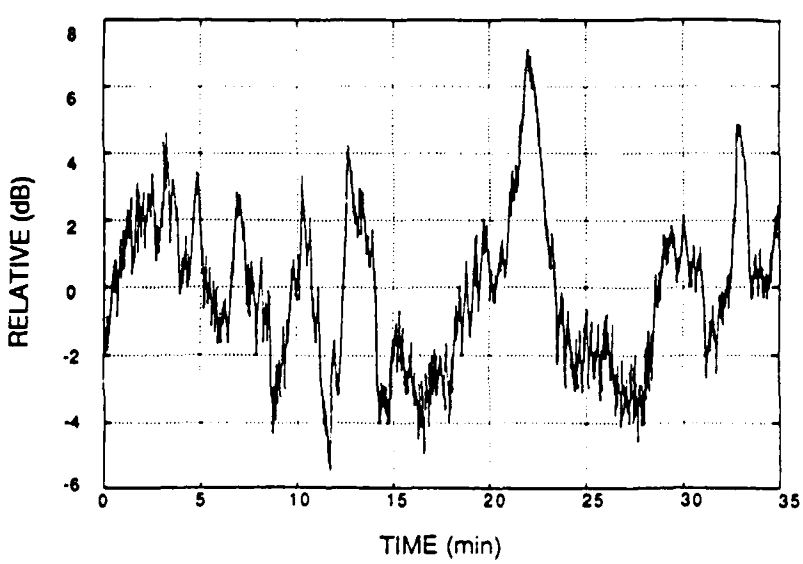


MITRE

VIEWGRAPH #17

RECEIVED POWER vs TIME: TYPICAL RESULT
Homestead, FL to Bedford, MA

March 20, 1989 from 1823Z



f_o = 25.885 MHz

Solarflux = 220

A_p = 20

σ = 2.27 dB

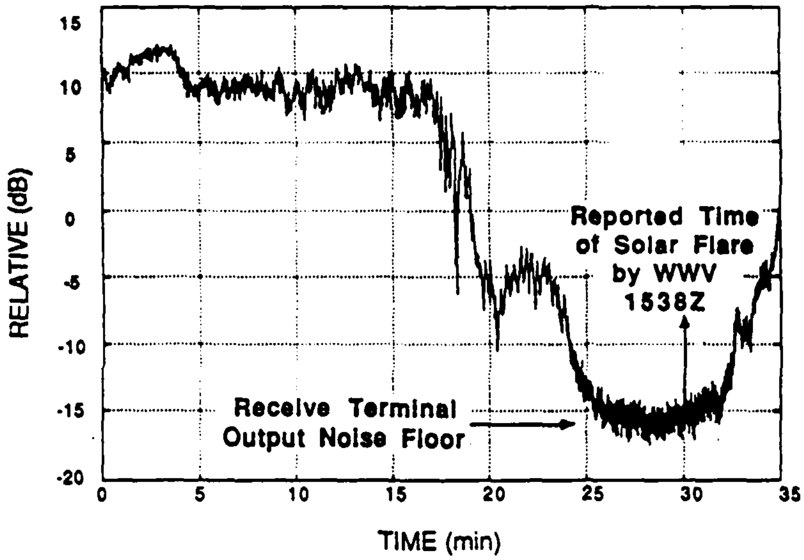
Decorrelation Time = 45 sec

MITRE

VIEWGRAPH #18

RECEIVED POWER vs TIME: ANOMALOUS RESULT
Homestead, FL to Bedford, MA

March 9, 1989 from 15088Z



f_o = 25.885 MHz

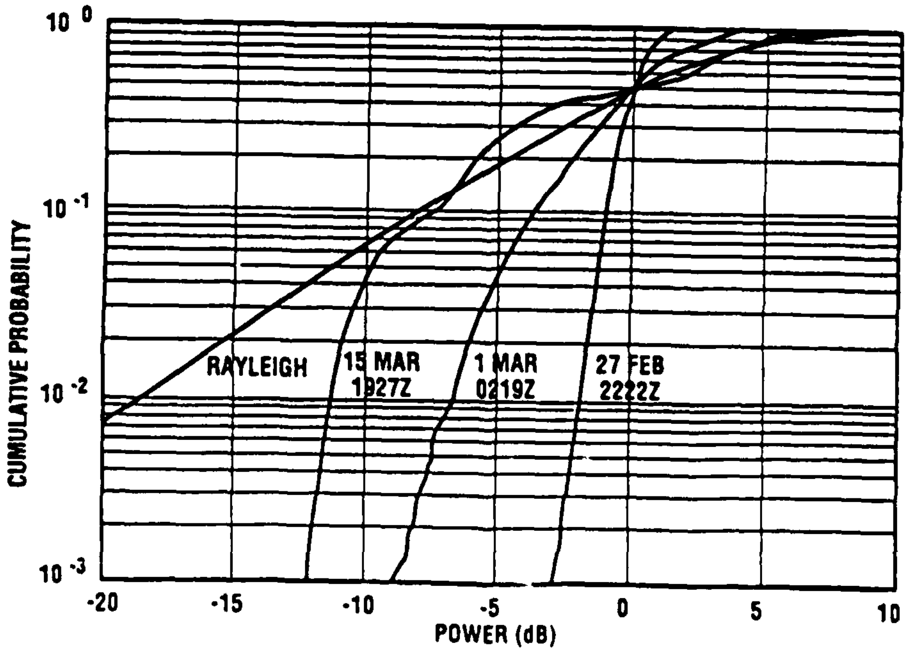
Solarflux = 207

A_p = 26

MITRE

VIEWGRAPH #19

CUMULATIVE PROBABILITIES OF RECEIVED POWER: THREE WIDEBAND HF CHANNELS AND A RAYLEIGH FADING CHANNEL



MITRE

VIEWGRAPH #20

MIDLATITUDE CHANNEL MEASUREMENTS SUMMARY

(2/27/89 -- 3/23/89 6-26 MHz, DAYTIME/NIGHTIME)

	Received Power vs Time Experiment 38 Runs - 35 min each		Doppler Shift Delay Spread Experiment 33 Runs	
	Standard Deviation (dB)	Decorrelation Time (sec)	Doppler Shift (Hz)	Delay Spread (µs)
lowest	0.71	5	0.00	2
highest	4.91	350	0.66	35
median	2.34	37	0.16	7

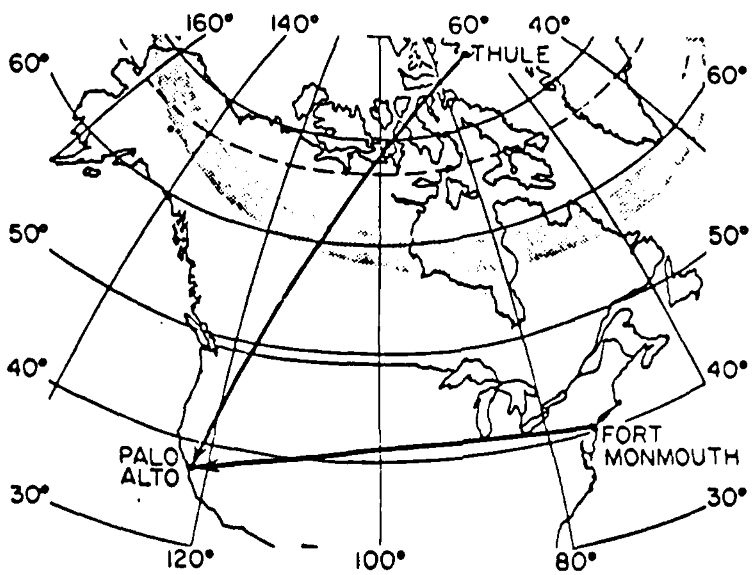
MITRE

VIEWGRAPH #21

MAP OF PROPAGATION PATHS

(CARRIER TRANSMISSION ONLY)

SRI MEASUREMENTS, 1964:
R. A. Shepherd and J. B. Lomax, "Frequency Spread in Ionospheric Radio Propagation", IEEE Trans. on Comm.Tech., April 1967



MITRE

VIEWGRAPH #22

measure the power spectrum of the received signals. What you see here [VIEWGRAPH #23] is Doppler shift as a function of time that they've measured. And you can see the size of it and how rapidly it varies. Truly the Doppler shift is varying slowly. It has excursions of plus or minus a Hertz roughly. They collected data over long periods of time and this is probably actual distribution of Doppler shifts, one of the few available in the world. These are the types of measurements that you need to predict performance of the rake modem. You can predict performance conditional on a given Doppler, conditional on the model delay dispersion, conditional on the given power, etc., but what are the probability distributions of these parameters? Here are empirical probability distributions of Doppler shift based upon the SRI measurements taken years ago. We're just starting to carry out some needed measurements again. Note that there are two traces on VIEWGRAPH #24, one for the Fort Monmouth path and one for Thule. The Thule is a disturbed channel. The Doppler shift has around the same statistics but you'll see the Doppler spread, which I'll discuss, is much worse for the disturbed channel.

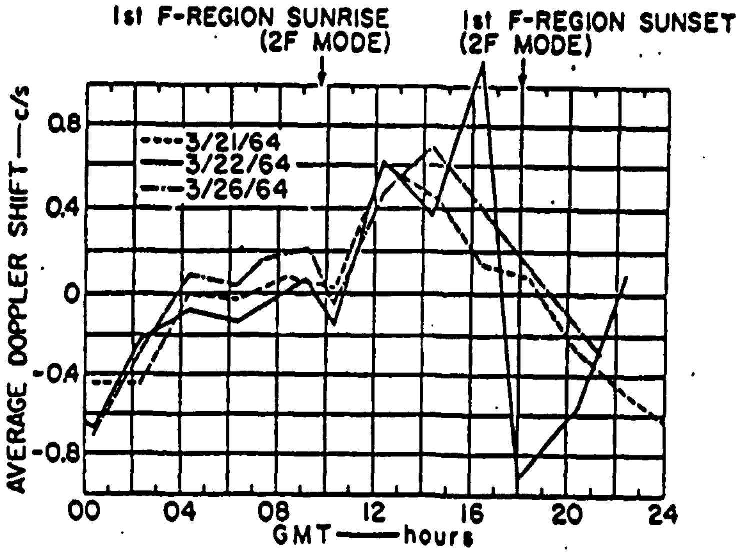
Here's the idealization of an ionogram of a disturbed channel [VIEWGRAPH #25] and here's a model for a disturbed channel [VIEWGRAPH #26]. You have a slow change in Doppler shift also but we replaced that simple dispersive time-invariant filter by a wide-sense stationary uncorrelated scattering channel, in which the tapped delay line model has randomly fluctuating independent complex Gaussian tap weights. The thing about that particular model is that if you are able to specify what's called the *scattering function* which describes the way power intensity is scattered in delay and Doppler, you have a complete statistical description. To be honest, you'd have to ask yourself, "Well, how good a model is that anyway?" And I wouldn't say that that's totally proven, but it's not bad from my point of view at this stage. We're collecting more information. The scattering function has had several names [VIEWGRAPH #27]; Price and Green in 1960 called it the "target scattering function." Years ago I called it the "scattering function." Leon Wittmer, at the Defense Nuclear Agency, has done a lot of work quite independently and reinvented it. He called it the "generalized power spectrum." And other people have called it the "delay-Doppler power spectrum," which is probably the most appropriate terminology.

This is one of my few slides with equations [VIEWGRAPH #28], just to motivate what this scattering function is. The complex Gaussian WSSUS channel can be viewed as sort of a densely tapped-delay line with infinitesimal independently fluctuating scatterers. Each scatterer has a time-varying complex gain which is the $g(t,\xi)d\xi$. And basically, it is the power spectrum of the fluctuation at the delay ξ which is the scattering function of delay-Doppler spectrum. I've just gone over it briefly for the lack of time, but you can just visualize it as being the power spectrum of the fluctuations at a given delay. You can relate this to the tapped delay line model (see VIEWGRAPH #29), but I'm not going to present it because there isn't enough time. So you visualize the scattering function as something like a two-dimensional power spectrum [VIEWGRAPH #30], and when you integrate over Doppler, you get what is called "delay power spectrum" which gives the multipath profile. The width of that delay profile is the delay spread or the multipath spread. On the other hand, if you inte-

AVERAGE DOPPLER SHIFTS

Fort Monmouth to Palo Alto Path, March 1964

SRI MEASUREMENTS, 1964:
R. A. Shepherd and J. B. Lomax, "Frequency Spread in Ionospheric Radio Propagation", IEEE Trans. on Comm.Tech., April 1967

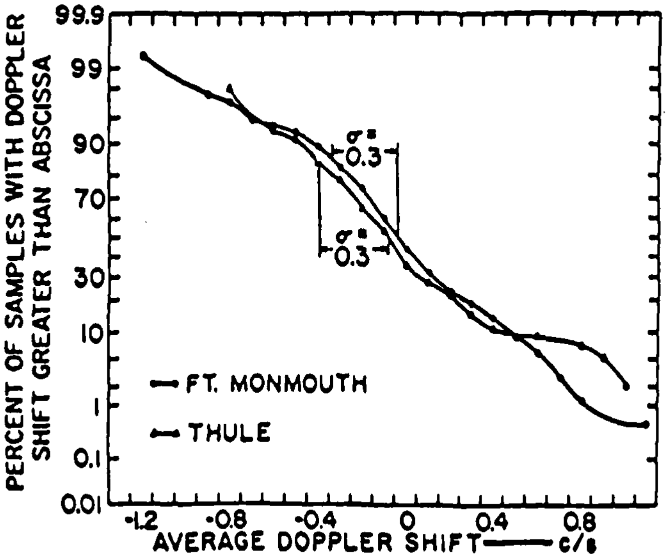


MITRE

VIEWGRAPH #23

DISTRIBUTION OF AVERAGE DOPPLER SHIFTS ON NORMAL PROBABILITY SCALE

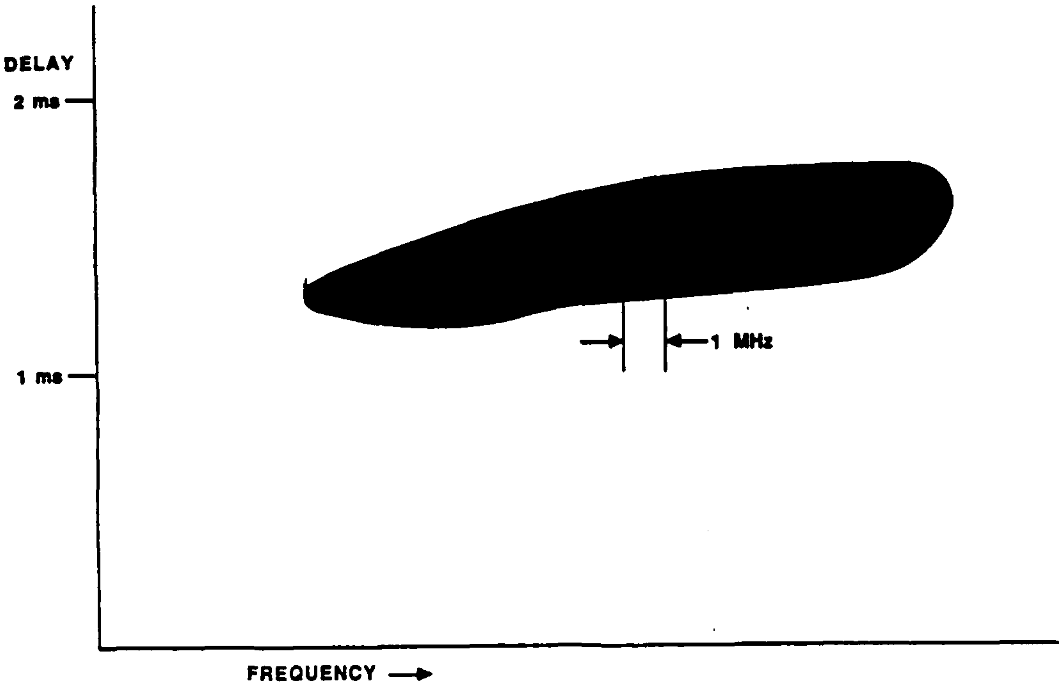
SRI MEASUREMENTS, 1964:
R. A. Shepherd and J. B. Lomax, "Frequency Spread in Ionospheric Radio Propagation", IEEE Trans. on Comm.Tech., April 1967



MITRE

VIEWGRAPH #24

IDEALIZED IONOGRAM EXAMPLE FOR DISTURBED CHANNEL

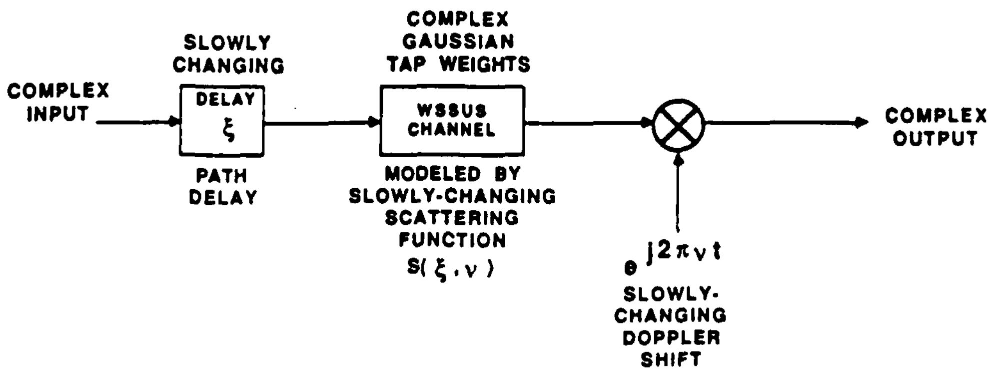


MITRE

VIEWGRAPH #25

QUASI-STATIONARY MODEL FOR DISTURBED 1 MHz WBHF PROPAGATION MODE

(USUALLY CONFINED TO TRANS-AURORAL, TRANS-EQUATORIAL, AND SOME EVENING MID-LATITUDE PATHS)



VIEWGRAPH #26

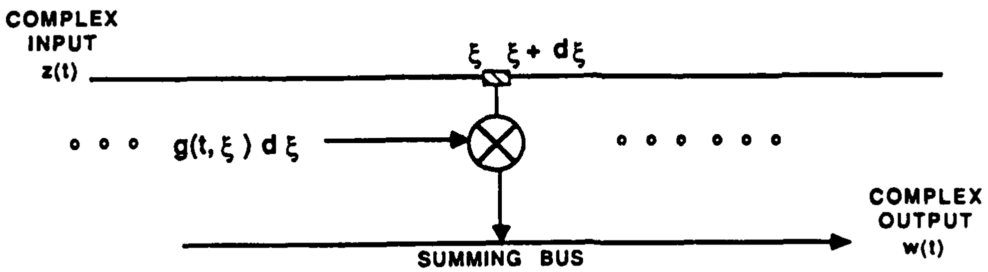
MITRE

SCATTERING FUNCTION NAMES

TARGET SCATTERING FUNCTION	R. Price & P. Green	1960
SCATTERING FUNCTION	P. Bello	1963
GENERALIZED POWER SPECTRUM	L. Wittwer	1979
DELAY DOPPLER SPECTRUM		

MITRE
VIEWGRAPH #27

COMPLEX GAUSSIAN WSSUS CHANNEL



$g(t,\xi)$ IS COMPLEX GAUSSIAN PROCESS FOR FIXED ξ

$$w(t) = \int z(t-\xi)\, g(t,\xi)\, d\xi$$

$$\overline{g^*(t,\xi)\, g(t+\tau,\eta)} = Q(\tau,\xi)\,\delta(\eta-\xi)$$

$Q(\tau,\xi)$ IS AUTOCORRELATION FUNCTION OF FLUCTUATIONS AT PATH DELAY ξ

$$S(\xi,\nu) = \int Q(\tau,\xi)\, e^{-j2\pi\nu\tau}\, d\tau \triangleq \text{DELAY-DOPPLER SPECTRUM}$$

$=$ POWER SPECTRUM OF FLUCTUATIONS AT DELAY ξ

MITRE
VIEWGRAPH #28

TAPPED DELAY LINE MODEL

$$g(t, \xi) = \sum g_k(t) \delta(\xi - k\Delta)$$

- COMPLEX TAP GAIN FUNCTIONS $g_k(t)$ BECOME COMPLEX GAUSSIAN PROCESSES
- THE CORRELATION FUNCTION $\overline{g_k^*(t)\, g_l(t+\tau)}$ DETERMINED BY $S(\xi, \nu)$ AND TERMINAL EQUIPMENT BANDLIMITING FILTER IMPULSE RESPONSE, $h(t)$.
- IF $S(\xi, \nu)$ CHANGES LITTLE FOR Δ CHANGES IN ξ THEN $\{g_k(t)\}$ BECOME STATISTICALLY INDEPENDENT AND POWER SPECTRUM OF $g_k(t)$ BECOMES

$$P_k(\nu) = \int |h(\xi - k\Delta)|^2 \, S(\xi, \nu) \, d\xi$$

$\sim S(\xi, \nu)$ IF $S(\xi, \nu)$ CHANGES LITTLE FOR DURATION OF $|h(\xi)|^2$

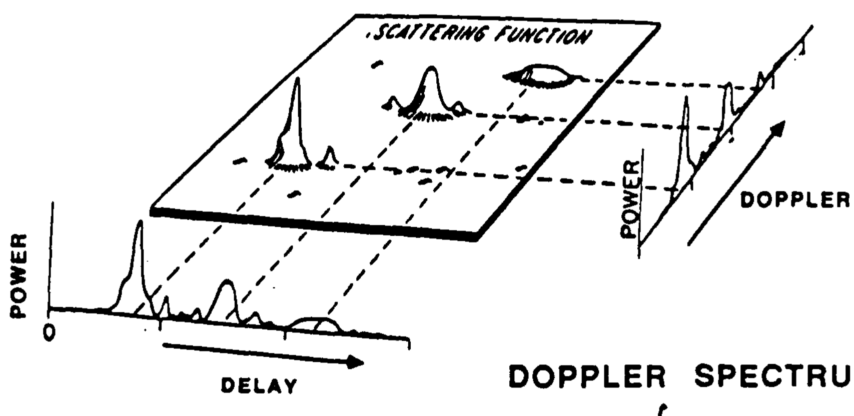
MITRE

VIEWGRAPH #29

DELAY SPECTRUM AND DOPPLER SPECTRUM

SCATTERING FUNCTION

$S(\xi, \nu)$



DOPPLER SPECTRUM

$$P(\nu) = \int S(\xi, \nu) \, d\xi$$

DELAY SPECTRUM

$$Q(\xi) = \int S(\xi, \nu) \, d\nu$$

VIEWGRAPH #30

MITRE

grate over delay, you get the Doppler power spectrum which is the power spectrum of the received carrier. The width of this power spectrum is called the Doppler spread.

Now we will talk about some measurements, typical Thule and Fort Monmouth to Palo Alto [VIEWGRAPH #31] measurements. This view shows the Doppler power spectrum, which is the received power spectrum corresponding to a transmitted carrier. Now you look at the power spectrum and you say, "My God, that thing is terribly jagged." I have to tell you this. After talking to the people who've done the measurements I discovered that instead of measuring power spectrum they measured periodograms. I know this audience knows the difference between a periodogram and a power spectrum ... I hope. The difference is if you take a random process and integrate over some finite time interval to compute a Fourier transform, you get a spectrum. If you form the magnitude squared, it's a periodogram, but it's not a power spectrum. You have to do further averaging. And that fine structure is due to the fact that they didn't average. They didn't average over frequency, or do successive snap shots. And I find out that everyone who has carried out these types of bound measurements have all measured periodograms. So when you look at them ... no, that's true, one didn't, I'll show you at the end. But it seems to be characteristic that they forget to do the averaging so it gets very ragged. But when they do the average, they come out much more reasonable. Now this [VIEWGRAPH #32] is interesting because this was over the SRI of the Fort Monmouth path which is supposed to be non-disturbed. It should have been well behaved.

Now let me go back to the previous one. [VIEWGRAPH #31] This is the Fort Monmouth-Palo Alto path. This has a very narrow spread formed by adding all the propagation modes together. You can see the difference between that and Thule-Palo Alto. However, this [VIEWGRAPH #32] is from SRI to Fort Monmouth and it's got a big mess. It's shifted, it has a mean Doppler shift which is off to one side and totally uncharacteristic, and it turns out it is due to some of the Aurora, the plasma, drifting off and providing off-axis reflections because there's a wide beam antenna.

Doppler spread is the width of the power empirical cumulative probability distributions for the spectrum of the fading. And this [VIEWGRAPH #33] shows the difference between Fort Monmouth and Thule-Palo Alto. Rather, note there's a ten to one difference in Doppler spread for the two paths.

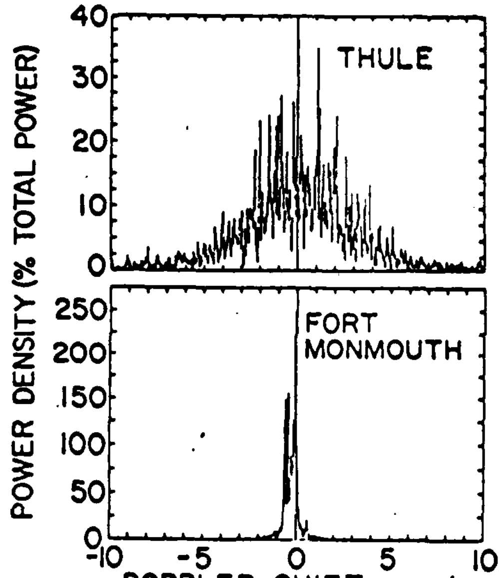
Len Wagner of NRL has taken some scattering function measurements [VIEWGRAPH #34]. You see how striated it is, he did not do the averaging and that's the result of doing the periodogram instead of a power spectrum. Now here are some SRI measurements (see VIEWGRAPHS #35,37,39) taken more recently in Greenland for the disturbed channel. And those are 5 dB contours for the scattering function. They are much smoother because he averaged his periodogram in the frequency domain. But when he did, he said it was for cosmetic purposes because it was too wild. He didn't realize that he was doing the right thing.

Here's some theoretical predictions. [VIEWGRAPHS #36,38,40] There's a fellow over at Mission Research Corp. working on a contract from DNA, a fellow by the name of Nickish who has done some very interesting work. When SRI did these most recent experiments, they had a radar, a back scatter radar and they got some indirect data on elec-

TYPICAL POWER SPECTRA

Thule, Greenland and Fort Monmouth-to-Palo Alto Paths

SRI MEASUREMENTS, 1964:
R. A. Shepherd and J. B. Lomax, "Frequency Spread in Ionospheric Radio Propagation", IEEE Trans. on Comm.Tech., April 1967



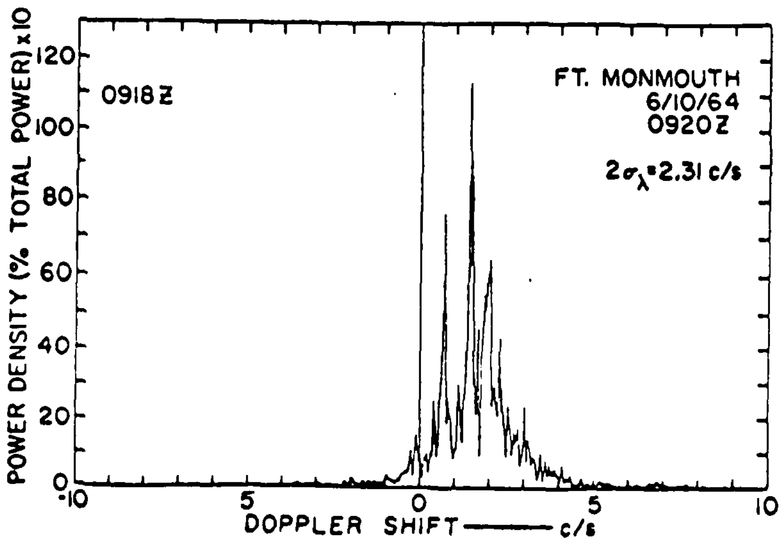
MITRE

VIEWGRAPH #31

AN OFF-PATH MODE

Fort Monmouth to Palo Alto Path, March 1964

SRI MEASUREMENTS, 1964:
R. A. Shepherd and J. B. Lomax, "Frequency Spread in Ionospheric Radio Propagation", IEEE Trans. on Comm.Tech., April 1967

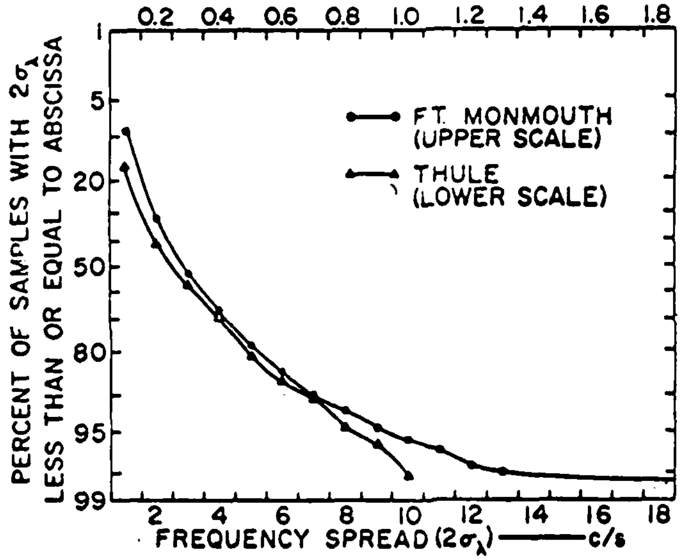


MITRE

VIEWGRAPH #32

DISTRIBUTIONS OF DOPPLER SPREAD ON NORMAL PROBABILITY SCALE

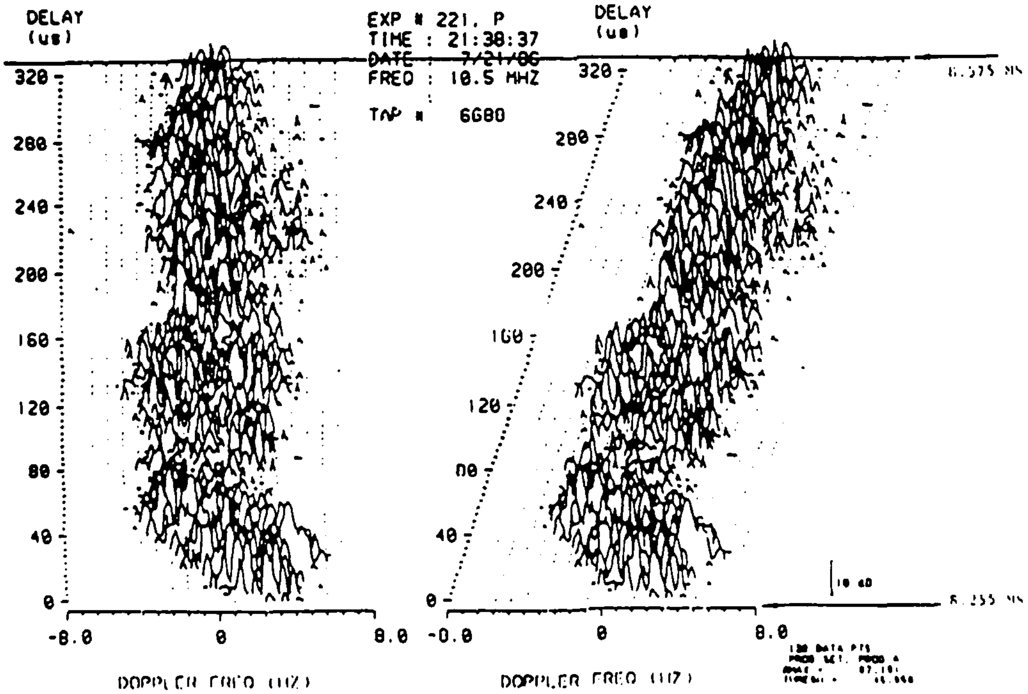
SRI MEASUREMENTS, 1964:
R. A. Shepherd and J. B. Lomax, "Frequency Spread in Ionospheric Radio Propagation", IEEE Trans. on Comm.Tech., April 1967



MITRE
VIEWGRAPH #33

EXAMPLE OF DISTURBED CHANNEL MEASUREMENTS
Frobisher Bay, Canada to Rome, New York

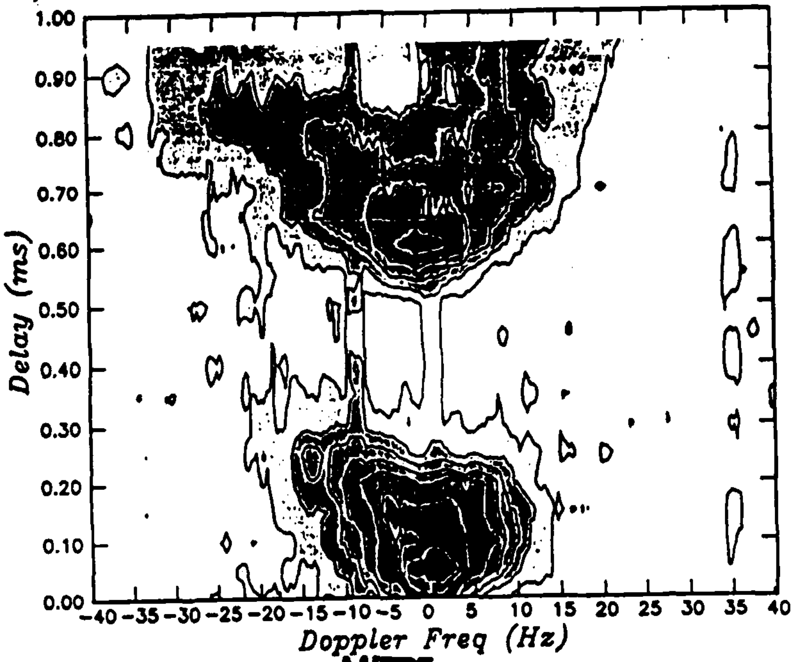
(from NRL Experiments by Len Wagner)



MITRE
VIEWGRAPH #34

HF CHANNEL PROBE
MEASURED SCATTERING FUNCTION

SRI MEASUREMENTS, Thule - Narssarssuaq, Greenland for March 20, 1985 14:18 UT (5 dB contours).
R. P. Basler et al., "HF Channel Probe", DNA-TR-85-247, May 1985.

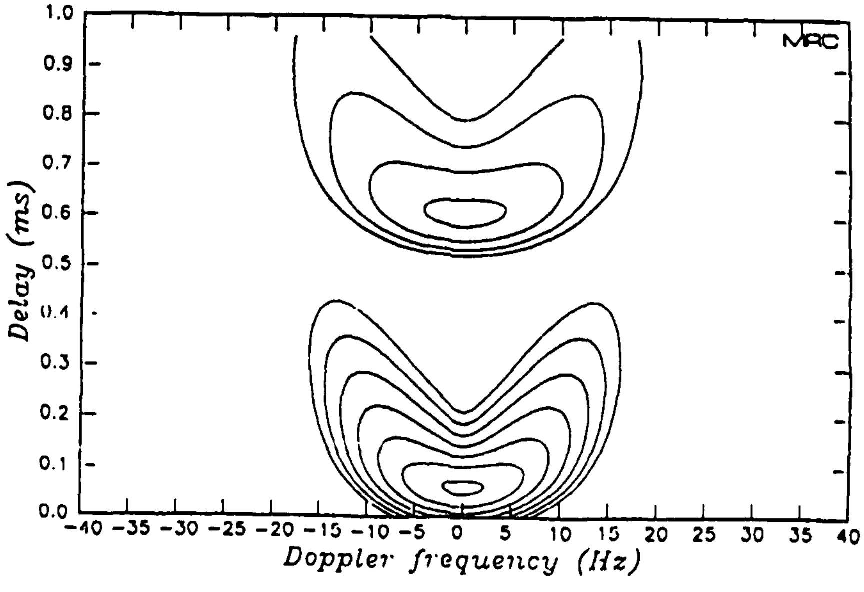


MITRE

VIEWGRAPH #35

SCATTERING FUNCTION FIT OF THE
MEASURED SCATTERING FUNCTION

SRI MEASUREMENTS, R. P. Basler et al. Thule - Narssarssuaq, Greenland for March 20, 1985 14:18 UT (5 dB contours).
ANALYSIS by L. J. Nicklsch, Mission Research Corporation, Carmel, California.
Report MRC/MRY-R-012, 31 December 1988.

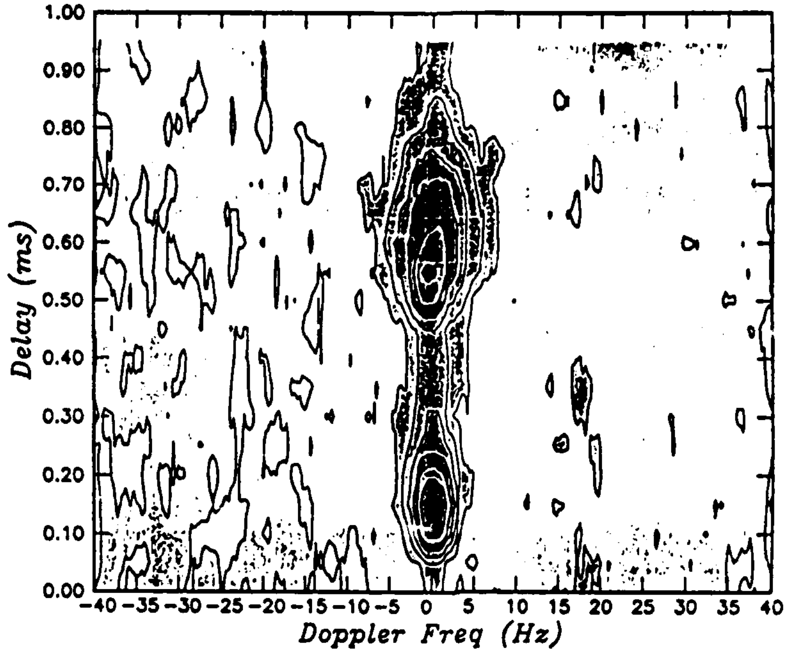


MITRE

VIEWGRAPH #36

HF CHANNEL PROBE MEASURED SCATTERING FUNCTION

SRI MEASUREMENTS, Thule - Narssarssuaq, Greenland for March 20, 1985 21:08 UT (5 dB contours).
R. P. Basler et al., "HF Channel Probe", DNA-TR-85-247, May 1985.

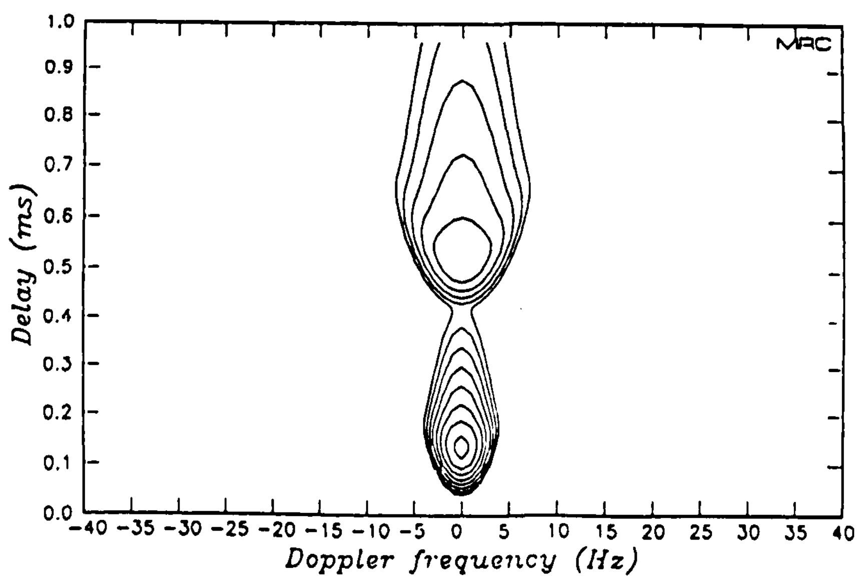


MITRE

VIEWGRAPH #37

SCATTERING FUNCTION FIT OF THE MEASURED SCATTERING FUNCTION

SRI MEASUREMENTS, R. P. Basler et al. Thule - Narssarssuaq, Greenland for March 20, 1985 21:08 UT (5 dB contours).
ANALYSIS by L. J. Nickisch, Mission Research Corporation, Carmel, California.
Report MRC/MRY-R-012, 31 December 1988.

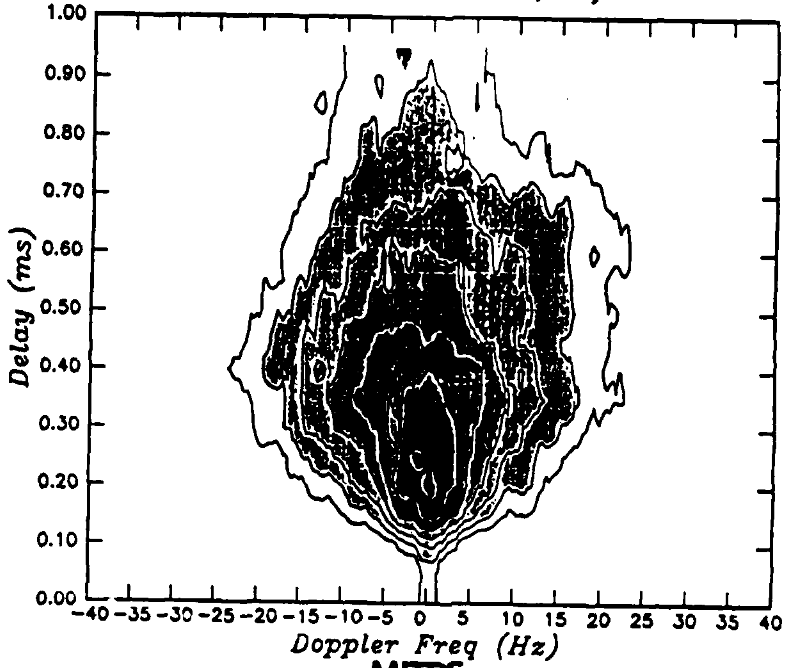


MITRE

VIEWGRAPH #38

HF CHANNEL PROBE
MEASURED SCATTERING FUNCTION

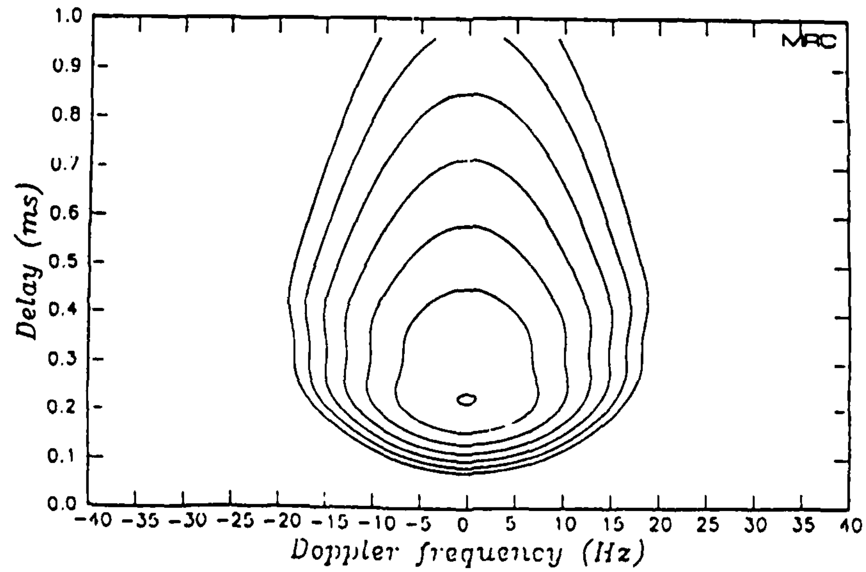
SRI MEASUREMENTS, Thule - Narssarssuaq, Greenland for October 17, 1984 22:03 UT (5 dB contours).
R. P. Basler et al., "HF Channel Probe", DNA-TR-85-247, May 1985.



MITRE
VIEWGRAPH #39

SCATTERING FUNCTION FIT OF THE
MEASURED SCATTERING FUNCTION

SRI MEASUREMENTS, R. P. Basler et al. Thule - Narssarssuaq, Greenland for October 17, 1984 22:03 UT (5 dB contours).
ANALYSIS by L. J. Nickisch, Mission Research Corporation, Carmel, California.
Report MRC/MRY-R-012, 31 December 1988.



MITRE
VIEWGRAPH #40

tron density fluctuation. He used this data, together with some theoretical work to predict what the scattering function would look like theoretically. He's done some pretty good modeling and I'll show you two more examples of the modeling. This is another case and here's his model. Here's another case and this is his model. But going back, look at the size of the spreads here. We see close to one millisecond of a delay spread, and the Doppler spreads are + or - 10 Hz. This is rapid fading, with significant amounts of multipath spread. So in summary, I say there's enough parameter information to allow the use of this quasi-stationary model, at least as a start for modem design which is an iterative process. We have to take more measurements. A lot more measurements are needed.

I wanted to say something about additive noise. The noise problem on the HF channel is far worse than the propagation channel, which is bad enough. You see [VIEWGRAPH #42], you have man made noise which varies very much geographically, you have atmospheric noise which varies seasonally, and you have interference. The interfering stations are a real mess. This shows measurements of a whole HF band. [VIEWGRAPH #43] You see all those large interfering areas there. The difference between the upper and lower figure is between day and night. You don't propagate well at high frequencies at night and everybody starts to crowd their transmissions down at lower frequencies. So the congestion gets even worse at night. This [VIEWGRAPH #44] is a 1 MHz power spectrum. We took an FFT of a 105 millisecond segment of a 1 MHz bandwidth portion of the HF channel, where we are, at Bedford in Massachusetts. Now you look at that and see all those spikes there, those are all interferers. You see how high they go. Some go up close to 127 dB as compared to the background noise level which is around 115 dB. Now we say you can communicate over that mess.

The next two slides [VIEWGRAPHS #45,46] say something about models developed. I need to go back one [VIEWGRAPH #44]. You could set a threshold and get cumulative distributions of power exceeding the threshold and try to develop an idea of how the power where the threshold varies with the threshold level. So, this was done experimentally and you get curves like this [VIEWGRAPH #45] for cumulative distribution. What it turns out, is after taking many of them and looking at data in Europe, you find out that the cumulative distributions can be modeled by a truncated Pareto distribution, which a log-log scale looks like [VIEWGRAPH #45]. But the important thing is if you study this viewgraph here, you ask the following important question: As you begin to eliminate parts of a band in order to remove the interference, how much of the band do you have to eliminate to cut down the interference to significant values? Because if you could just remove only 10-20% of the band and cut down the interference by 30 dB, you would have a chance of communicating. That's roughly what VIEWGRAPH #47 shows for that particular set of measurements taken at Bedford. You notice the axis at the bottom, percentage of the band excised. At 10% excision, looking upwards, the intersection, you see you're a little over 180 dB as opposed to 210 dBs for 0% excision. You thus achieve a 30 dB improvement with a 10% excision. Characteristically, looking at recent data, it appears you can achieve 20-30 dB reduction in the interference power with 10-20% of the band excised. Excising 10 or 20 % of the band will only reduce the power in your spread spectrum signal by 1 dB, but

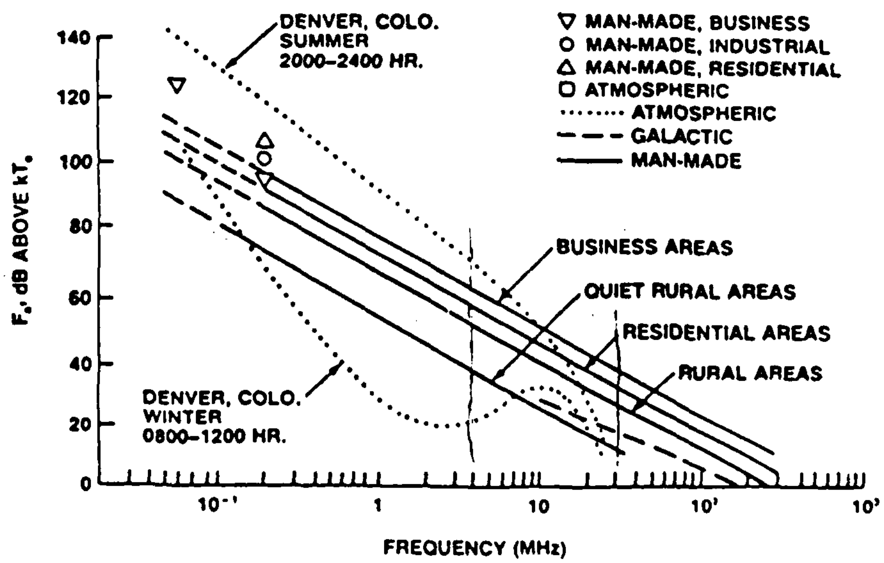
SUMMARY

- ENOUGH PARAMETER INFORMATION IS AVAILABLE TO ALLOW USE OF THE PROPOSED QUASI-STATIONARY MODELS AS A BASIS FOR EXPERIMENTAL MODEM DESIGN AND PERFORMANCE ANALYSIS.
- MUCH MORE QUASI-STATIONARY PARAMETER MEASUREMENTS ARE NEEDED TO PREDICT LONG-TERM PERFORMANCE.

MITRE

VIEWGRAPH #41

TYPICAL NOISE LEVELS



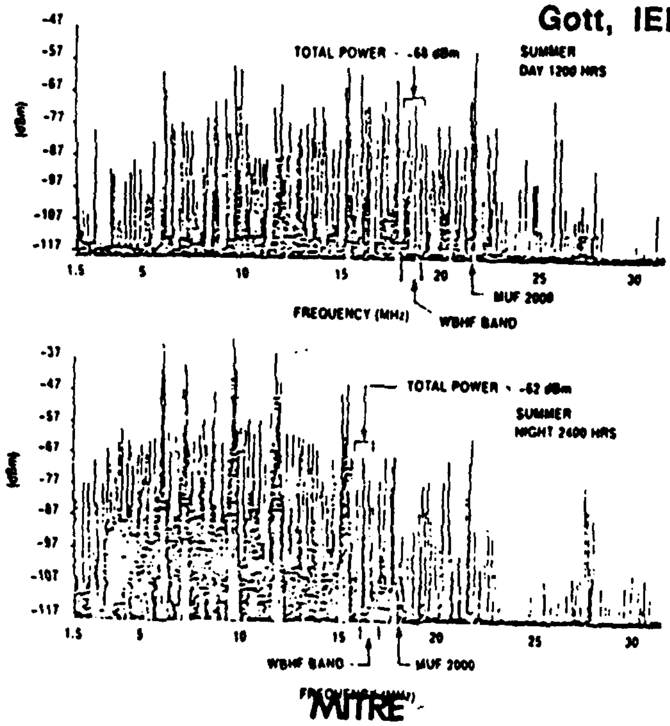
MEASURED AND EXPECTED NOISE LEVELS FOR DENVER, COLORADO (CCIR AND ITS MODELS)

MITRE

VIEWGRAPH #42

HF BAND SPECTRA
SUMMER SOLSTICE, CENTRAL ENGLAND

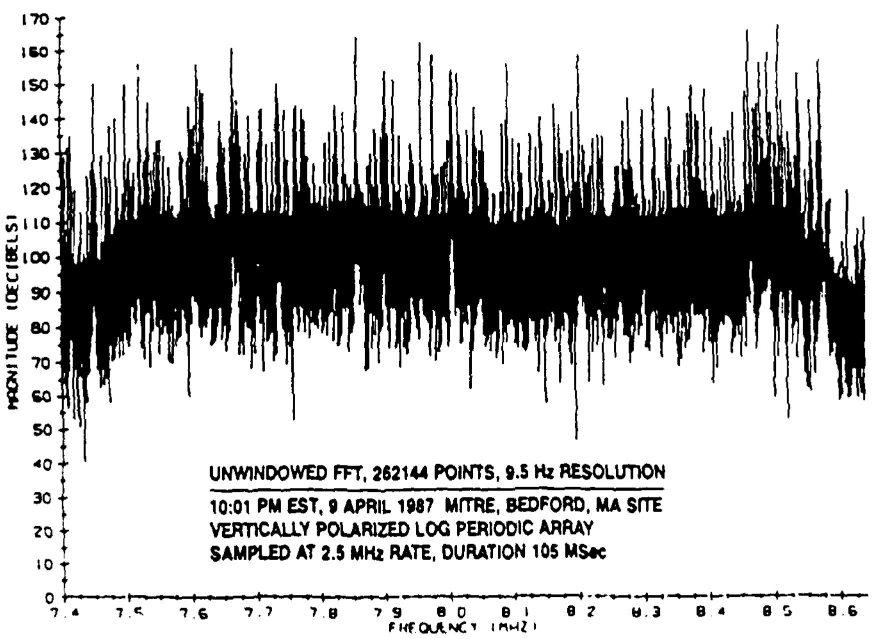
Gott, IEE Proc., Dec.1985



MITRE

VIEWGRAPH #43

WIDEBAND HF NOISE AND INTERFERENCE
Full 1 MHz Spectrum at 8.0 MHz



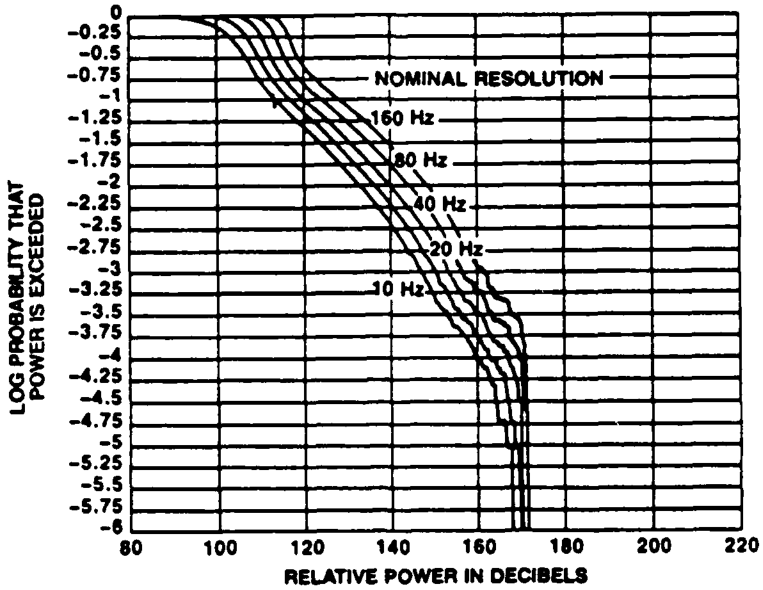
MITRE

VIEWGRAPH #44

WIDEBAND HF OCCUPANCY

Cumulative Probability Distributions, 1 MHz Bandwidth

B. D. Perry and L. G. Abraham, "Wideband HF Interference and Noise Model Based on Measured Data", M88-7, MITRE Corporation, March 1988. IEE Conference Publication 284, 4th International Conference on HF Radio Systems and Techniques, London, U.K., April 1988.



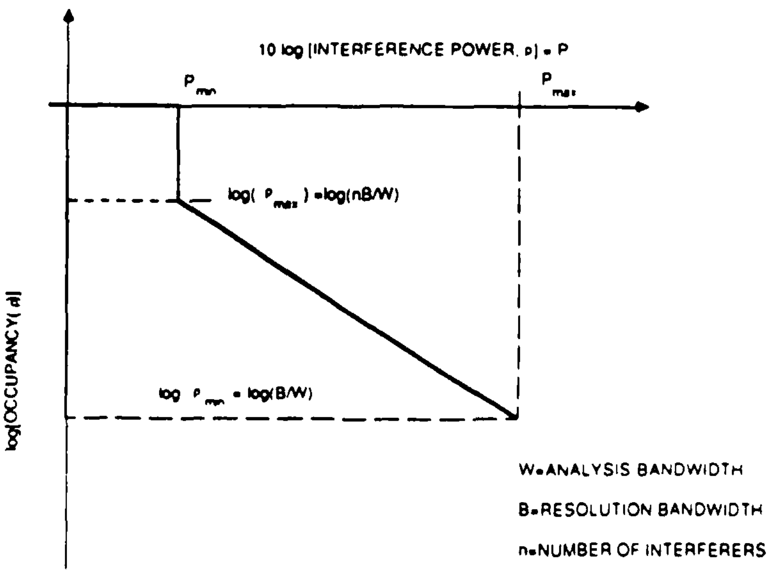
DURATION 105 msec, ALMOST 105,000 POINTS
10:01 PM EST, 9 APRIL 1987, MITRE/BEDFORD, MA SITE
VERTICALLY POLARIZED LOG PERIODIC ARRAY

MITRE

VIEWGRAPH #45

MODEL OF INTERFERENCE DISTRIBUTION

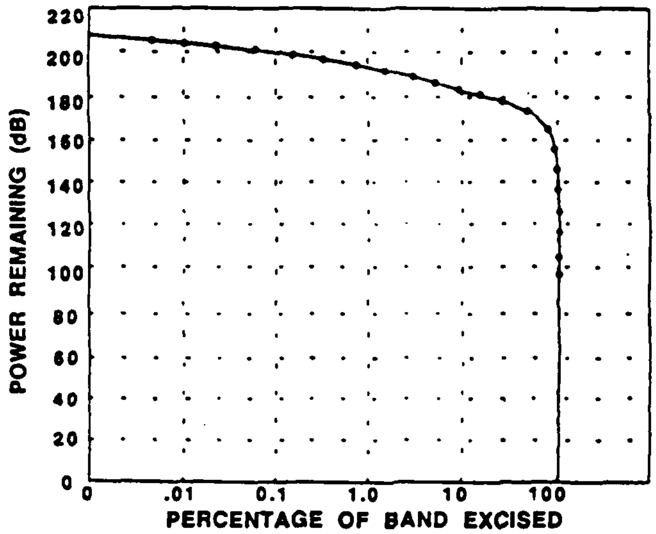
B. D. Perry and R. Rifkin, "Interference and Wideband HF Communications", Proceedings Fifth Ionospheric Effects Symposium, Springfield, VA, USA, 1987.



MITRE

VIEWGRAPH #46

POWER REMAINING AFTER EXCISION VS PERCENTAGE OF BAND EXCISED

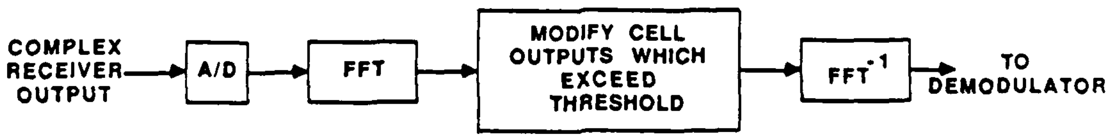


8 MHz CENTER FREQUENCY
38.14 Hz RESOLUTION 1 MHz BANDWIDTH
10:01 PM EST, 9 APRIL 1987, MITRE BEDFORD

MITRE

VIEWGRAPH #47

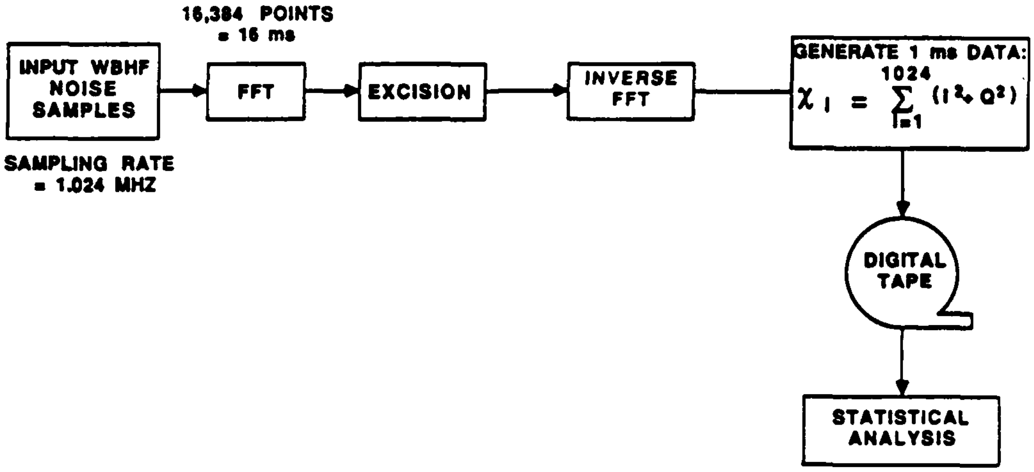
BLOCK FFT INTERFERENCE SUPPRESSION



VIEWGRAPH #48

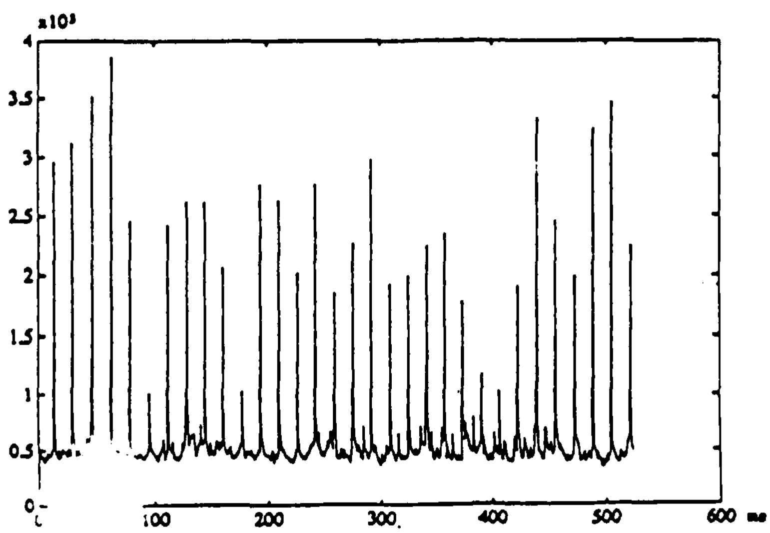
MITRE

1 ms POST-EXCISOR NOISE POWER MEASUREMENTS



MITRE
VIEWGRAPH #49

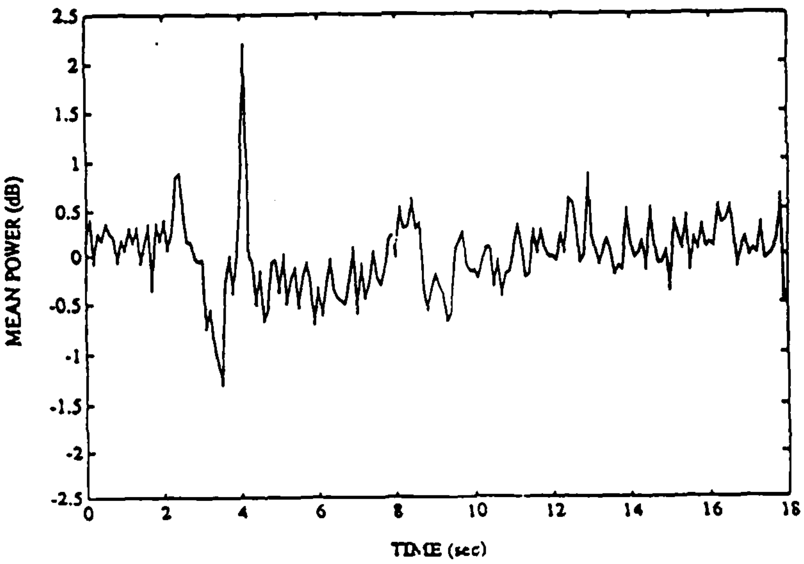
1 ms averages showing noise spike phenomenon occurring every 16 msec.



$$\chi = \sum_{i=1}^{1000} (I^2 + Q^2)_i$$

MITRE
VIEWGRAPH #50

Post-Excisor Residual Short Term Power vs Time for Averaging Time T = 96 msec

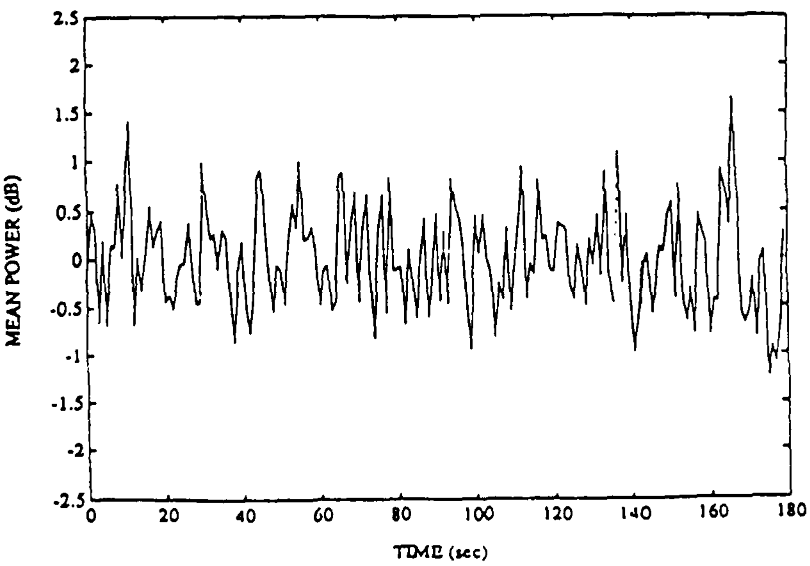


(25 October 1988, 14:00 G.M.T.)

MITRE

VIEWGRAPH #51

Post-Excisor Residual Short Term Power vs Time for Averaging Time T = 1 sec

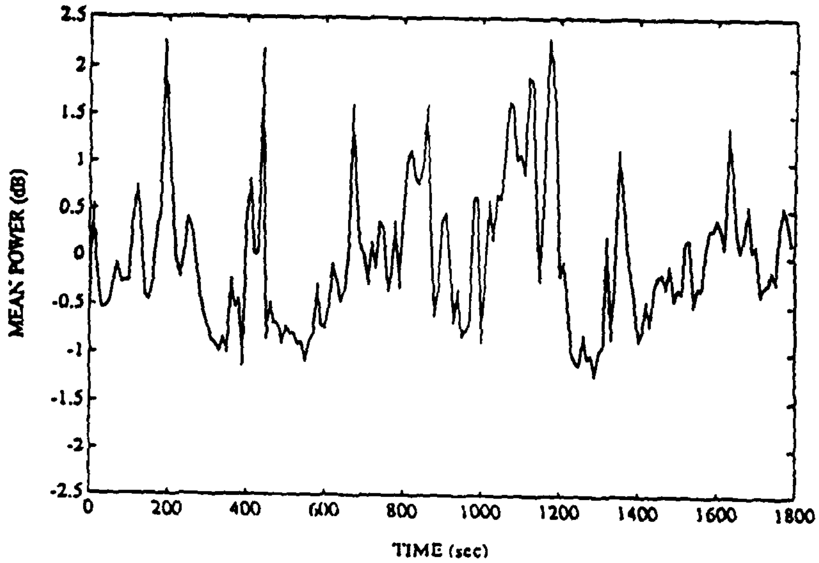


(25 October 1988, 14:00 G.M.T.)

VIEWGRAPH #52

MITRE

Post-Excisor Residual Short Term Power vs Time for Averaging Time T = 10 sec

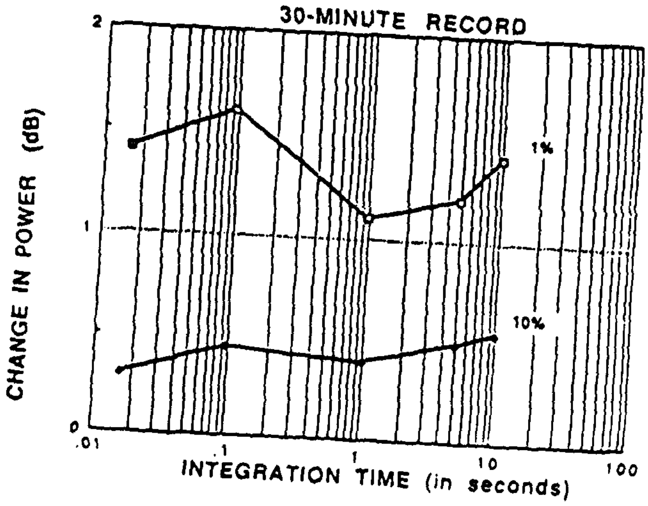


(25 October 1988, 14:00 G.M.T.)

MITRE

VIEWGRAPH #53

PERCENTILES FOR THE CHANGE IN AVERAGE POWER BETWEEN SUCCESSIVE INTEGRATION TIMES FOR POST-EXCISOR NOISE



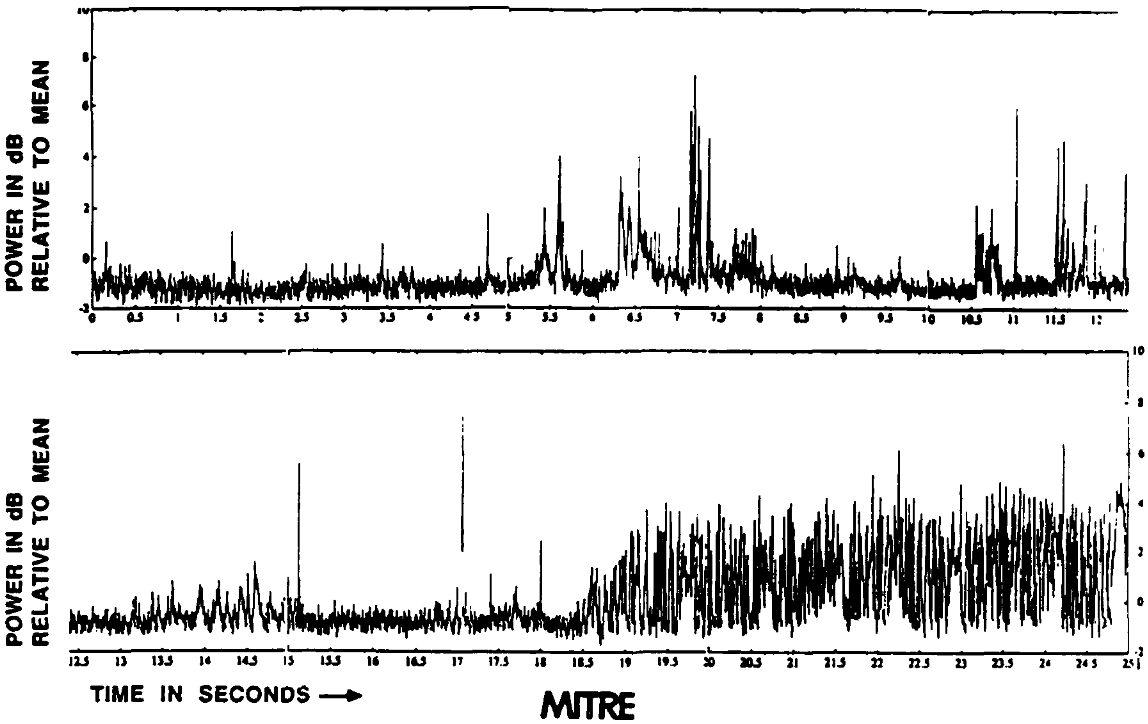
(9 January 1989, 12:22 G.M.T.)

VIEWGRAPH #54

MITRE

SAMPLE OF NOISE RECORD ILLUSTRATING NOISE BURSTS

(#TH052637. 12 January 1989. 12:26 AM E.S.T.)



MITRE

VIEWGRAPH #55

SUMMARY

- EXCISION PROCEDURE SHOULD PRODUCE SIGNIFICANT REDUCTIONS IN WBHF NOISE POWER WITH MODEST REDUCTIONS IN SIGNAL POWER.
- POST-EXCISION NOISE STILL NONSTATIONARY.
- WE'VE GOT A LONG WAY TO GO.

VIEWGRAPH #56

MITRE

WIDEBAND HF CHANNEL MODELING FOR DIRECT SEQUENCE SPREAD SPECTRUM MODEM DESIGN

Phillip A. Bello

Disturbed Channel Propagation-Theoretic Works Which Evaluate Channel Parameters

Henry G. Booker and Tao Jing-Wei, "A Scintillation Theory of the Fading of HF Waves Returned from the F Region: Receiver Near Transmitter," *Journal of Atmospheric and Terrestrial Physics* **49,** No. 9 (1987), 915–938.

Henry G. Booker, Jing-Wei Tao, and Amir B. Behroozi-Toosi, "A Scintillation Theory of Fading in Long Distance HF Ionospheric Communications." *Journal of Atmospheric and Terrestrial Physics* **49,** No. 9 (1987), 939–958.

J.-F. Wagen and K. C. Yeh. "Simulation of HF Propagation and Angle of Arrival in a Turbulent Ionosphere," *Radio Science* Vol. **24 No. 2** (March-April 1989), pages 196–208.

L. J. Nickisch, "The Mutual Coherence Function in Extended Moving Random Media," *MRC/MRY-R-012.Mission Research Corporation, Carmel, CA* (December 1988).

Wide Band HF Channel Measurements

L. S. Wagner, J. A. Goldstein and E. A. Chapman, "Wideband HF Channel Prober: System Report," *NRL Report No. 8622, Naval Research Labs, Washington, DC* **AD-A127-040** (1983).

L. S. Wagner and J. A. Goldstein, "High Resolution Probing of the HF Skywave Channel: F Layer Mode Fluctuations," *Effect of the Ionosphere on C^3I Systems, J. M. Goodman, editor* **Ionospheric Effects Symposium, L.C. 85-600558** (1984), 63–73.

L. S. Wagner and J. A. Goldstein, "High Resolution Probing of the HF Ionospheric Skywave Channel: F2 Layer Results," *Radio Science* **Vol.20 No. 3** (1985), 287–302.

Roy P. Basler, et al., "Ionospheric Distortion of HF Signals," *Radio Science* **Vol.23, No.4** (July–August 1988).N. 569–579..

will cut down the interference greatly. What we're using at MITRE, at the moment, is a quick and dirty exciser; take an FFT, a block at a time, modify those FFT bins that exceed the threshold by throwing them away or clipping them, then do an inverse FFT. Do this continually. Then you feed that to the rake demodulator. I don't think I'll discuss any more because I've run out of time.

LINDSEY: O.K. Phil, thank you very much. We'll reserve questions and comments till after our break, until all speakers have completed their discussions. At this point I would like to introduce Dr. Ken Wilson from Martin Marietta who will talk to us about transmission of information via of the optical channel, and ... O.K. J think I'll just shove this back. Leave it here ....

KEN WILSON: *Optical Channels*

[ST CTF vg1] Back in the 1984 time frame, Martin Marietta had to come up with a free-space optical communication link for inter-satellite use. At that time, Martin had the assistance of McDonald Douglas, one of the leaders in building such things. Suffice it to say that there was a falling out between contractors (a fact that plagues the relationship to this day), and Martin cancelled the contract with MacDac. Due to the tremendous inertia in a design by the time a program reaches the Preliminary Design Review stage, the method of coupling the laser diode output to the focus of the transmitting optical system was already fixed - a multimode step index fiber optic cable was used. We stayed with this design for some very good engineering reasons: it keeps the focal point constant, elimates any variation of the optical beam alignment with respect to thermal disturbances of the laser diode heat sink, keeps the laser diode heat off of the optical bench, eases optical alignment, and so on.

[ST CTF vg2] However, in staying with the fiber optic cable a problem is generated for the problem that is solved - you end up with speckle. It turned out that not only do you end up with spatial speckle that is static, but you end up with impacts on the waveform that you produce at the far end of the link, at the receiver. Because of the fact that the laser diode is shifting frequencies as the current is modulated, the spectrum at the input of the optical fiber is shifting, and the modal noise of the fiber gives a dynamic speckle problem. This produces waveform amplitude distortions at locations in the far field, causing a mismatch between the waveform and the matched filter with an attendant decrease in performance. This talk is only going to address the static portions of the speckle measurements that were made in an attempt to quantify this channel. Walt Bremmer also presented at MILCOM '88 a series of measurements (paper 32.6) that were dynamic in nature and that unfortunately were not a complete data set, that is, measurements of dynamic speckle as a function of position throughout the whole beam. We will show you complete data sets.

The problem faced by the communication system engineer is that of evaluating the effect of the dynamic speckle on the performance that can be delivered through the speckle channel. Therefore we seek a methodology by which the performance of the channel can be evaluated. One of the most inclusive channel models which has been formulated is Lindsey's Space-Time Communication Transmittance Function (ST CTF) channel model. The objective of this talk is to show the progress that has been made in the application of this very general model to the laser speckle channel.

[ST CTF vg3] I'm going to rush through

APPLICATION OF THE SPACE-TIME COMMUNICATION TRANSMITTANCE FUNCTION CHANNEL MODEL TO A FIBER OPTIC COUPLED SEMICONDUCTOR LASER DIODE

MAY 14, 1989

Dr. Kenneth E. Wilson
Mark A Hennecken

MARTIN MARIETTA
ASTRONAUTICS GROUP

ST CTF vg1

BACKGROUND

PROBLEM

- FIBER OPTIC COUPLING OF LASER DIODES USING MULTIMODE FIBERS GENERATES SPECKLE IN THE FAR FIELD. THE SPECKLE HAS SPATIAL AND TEMPORAL CHARACTERISTICS THAT DEFINE A SPACE-TIME CHANNEL WHEN PROJECTED FOR FREE SPACE LASER COMMUNICATION TRANSMISSION.

- THE IMPACT OF THIS DYNAMIC SPECKLE ON THE COMMUNICATION CHANNEL PERFORMANCE SHOULD BE EVALUATED ANALYTICALLY.

OBJECTIVE

- APPLY THE SPACE-TIME COMMUNICATION TRANSMITTANCE FUNCTION (ST CTF) CHANNEL MODEL TO THE LASER SPECKLE CHANNEL BY EVALUATING DATA THAT HAS BEEN TAKEN IN TERMS OF THE MODEL. THAT IS, MAKE A START AT BUILDING A CHANNEL MODEL.

ST CTF vg2

THE ST CTF CHANNEL MODEL

- THE MODEL AS ORIGINALLY FORMULATED BY LINDSEY AND LO RELATES THE OBSERVED ELECTROMAGNETIC FIELD TO THE FIELD THAT WOULD BE PRESENT IN THE ABSENCE OF THE MEDIUM. THE SPACE-TIME CHANNEL TRANSMITTANCE FUNCTION DEFINES THAT RELATIONSHIP:

$$\vec{E}_0(\vec{r}_0, \vec{r}_s; t) = \mathcal{H}^*(\vec{r}_0, \vec{r}_s; t)\ \vec{E}(\vec{r}_0, \vec{r}_s; t)$$

Medium Present CTF Medium Absent

- THIS CAN BE PUT IN MATRIX NOTATION

$$\begin{bmatrix} E_{ox} \\ E_{oy} \\ E_{oz} \end{bmatrix} = \begin{bmatrix} h_{xx} & h_{xy} & h_{xz} \\ h_{yx} & h_{yy} & h_{yz} \\ h_{zx} & h_{zy} & h_{zz} \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix}$$

ST CTF vg3

THE ST CTF CHANNEL MODEL (Continued)

- THE SPACETIME CORRELATION FUNCTION IS DEFINED IN TERMS OF THE h's:

$$R_{h_{ij}}(\vec{r}_1, \vec{r}_2, \vec{r}_s; t+\tau) = R_{h_{ij}}(\Delta\vec{r}, \vec{r}_s; \tau)$$

$$= \langle h_{ij}(\vec{r}_1, \vec{r}_s; t)\ h^*_{ij}(\vec{r}_2, \vec{r}_s; t+\tau) \rangle$$

- DEFINE THE NORMALIZED ST CTF CORRELATION FUNCTION AS

$$\rho_{h_{ij}}(\Delta\vec{r}, \vec{r}_s; t) = \frac{R_{h_{ij}}(\Delta\vec{r}, \vec{r}_s; t)}{R_{h_{ij}}(\vec{0}, \vec{r}_s; t)}$$

ST CTF vg 4

these because I assume that you are already knowledgable about Lindsey's ST CTF model. As originally formulated, it was a polarization-dependent model where the output is related to the input by the polarization matrix $H*$. The polarization matrix expresses the observed field (here on the left) in terms of the field that would exist at the observation point if the dispersive medium were absent (here on the right). For notational convenience a matrix can be used to represent the components of the vector field.

[ST CTF vg4] The correlation function is defined in terms of the components of $H*$. It is the expectation value of $H \times H*$ for the i^{th} and j^{th} components. i and j typically represent the spatial coordinates of the correlation function. We define a normalized space-time correlation function as the ratio of these correlations, normalized by the no-dispersion case at the same point. This ρ will figure in the integrals that we used to derive correlation time and correlation lengths.

[ST CTF vg5] And these are Bill's definition for correlation time and length. What is different about what we've done here is that Bill's model was formulated to handle polarization details of the channel. Most of the fluctuations we're going to see are caused by the polarization details of the medium. However, ...

[ST CTF vg6] Those details are essential in the radio-frequency channel. For this particular channel, we have a non-polarization sensitive receiver. We have a photodiode detector that if a photon hits it, produces a photoelectron with the probability associated with the quantum efficiency of the photo detecting material at that photon wavelength. The detector is polarization insensitive, therefore in terms of the space-time correlation function, the H_{ij} are diagonal. Furthermore, we have assumed that the optical axis is in the direction of the positive z axis, so there is no z-variation and H_{zz} in the matrix is equal to 1. The measurements were made in the steady state and incoherently, so the t component goes away. When the problem is considered in rigorous detail, you find that in order to actually model the speckle pattern that you produce in the far field, all of these effects have to be taken into account inside the optical fiber. But the beauty of using the channel model is that it permits you to ignore some of the physical details that you cannot actually compute, and we cannot compute the speckle pattern at the fiber output given the multimode, multi-spectral optical input provided by the laser diode source.

[ST CTF vg7] Right on into the source, we have a six-stripe laser diode with an output aperture of 56 microns driving a 75 micron diameter core optical fiber as shown on the next viewgraph. Let me go ahead and put it up and talk to the physical configuration directly.

[ST CTF vg8] Here is a six-stripe diode. This is about a 10-inch piece of step index optical fiber cable with the indices of refraction shown, having a 75 micron core and a 25 micron thick cladding. The input to the fiber is vertical because these particular diodes are 99% pure in polarization. The fiber is driven by multiple sources, the 6 stripes of the laser, each source having multiple spectral lines. I did not show you the output spectrum of this, let me see if I brought it with me [Yes, I did, I'll show it to you in a minute]. The laser is near field coupled: the laser/fiber separation distance is 10 to 20 microns as opposed to a source aperture of 56 microns, therefore you are in the near field. The stripes themselves are alternating in phase with a 180-degree phase shift. Without the fiber present

THE ST CTF CHANNEL MODEL (Continued)

- THEN THE CHANNEL CAN BE DESCRIBED IN TERMS OF CORRELATION LENGTHS AND CORRELATION TIMES:

$$\tau_c = \int_{-\infty}^{\infty} \rho_{h_{ij}}(\Delta \vec{r} = \vec{0}, \vec{r_s}; t) \; d\tau$$

$$L_x = \int_{-\infty}^{\infty} \rho_{h_{ij}}(\Delta x, \Delta y = \Delta z = 0, \vec{r_s}; 0) \, d\Delta x$$

$$L_y = \int_{-\infty}^{\infty} \rho_{h_{ij}}(\Delta x = 0, \Delta y, \Delta z = 0, \vec{r_s}; 0) \, d\Delta y$$

$$L_z = \int_{-\infty}^{\infty} \rho_{h_{ij}}(\Delta x = \Delta y = 0, \Delta z, \vec{r_s}; 0) \, d\Delta z$$

ST CTF vg5

THE ST CTF MODEL APPLICATION

- THE ST CTF MODEL WAS FORMULATED TO HANDLE THE POLARIZATION DETAILS OF THE CHANNEL. THESE DETAILS ARE ESSENTIAL TO THE RADIO FREQUENCY CHANNEL AND TO THE HETERODYNE LASER CHANNEL. THEY ARE ALSO IMPORTANT IN THE FIBER OPTIC COUPLED LASER CHANNEL IN COMPUTING THE SPECKLE PATTERN.
- NO SERIOUS ATTEMPT WAS MADE TO COMPUTE THE SPECKLE PATTERN. INSTEAD, THE PATTERN WAS MEASURED AND IS STUDIED STATISTICALLY.
- THE DETECTOR WE USE IS POLARIZATION INSENSITIVE.
- THE MEASUREMENTS WE MADE ARE POLARIZATION INSENSITIVE.
- THEREFORE, IN TERMS OF THE ST CTF, THE h_{ij} ARE DIAGONAL.
- THE MEASUREMENTS HAVE NO z VARIATION, SO $h_{zz} = 1$
- THE MEASUREMENTS WE MADE ARE STEADY STATE, SO $\tau = 0$.

ST CTF vg6

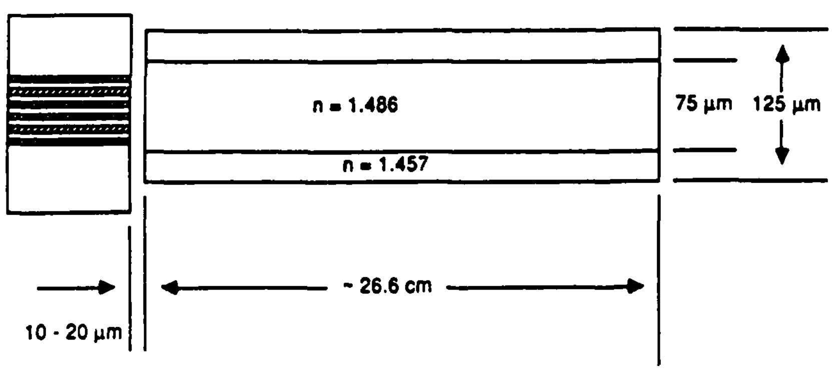
FIBER OPTIC COUPLED LASER DIODE SOURCE

- A 6 STRIPE DIODE WAS USED TO DRIVE A 75 µm OPTICAL FIBER AS SHOWN.
- THE STRIPES ARE PHASE LOCKED AT SEVERAL FREQUENCIES (WAVELENGTHS) BY EVANESCENT COUPLING BETWEEN STRIPES.
- THE DIODE IS NEAR-FIELD COUPLED TO THE FIBER. THIS RESULTS IN A 'SEVERAL COHERENT SOURCE' INPUT EXCITATION OF THE "MODES" OF THE FIBER. THE VARIOUS PATH LENGTHS (DIFFERENT FOR EACH MODE) CAUSE CONSTRUCTIVE AND DESTRUCTIVE INTERFERENCE AT THE FIBER OUTPUT, PRODUCING THE INTENSITY VARIATION WE CALL SPECKLE.
- THE TRANSMITTING LENS IMAGES THE FIBER OUTPUT, PROJECTING IT TO A SPACE LIKE HYPERSURFACE AT ∞.

ST CTF vg7

LASER SOURCE

6 Stripe diode



Laser Input to Fiber:
Polarization: vertical
Multiple Sources (6)
Multiple Spectral Lines
Near Field Coupled
Alternating 180° Phase Shift
Each stripe has gain at several distinct wavelengths
Each wavelength is phase locked across the stripes
Individual stripe divergence ~ 20° x 40°

Fiber:
Spectran 75
75 ˜m core, n = 1.486
125 µm cladding dia, n = 1.457
Step Index
Multimode
Non-polarization preserving
Length ~ 10.5 in ~ 26.6 cm
Numerical Aperture = 0.292

ST CTF vg8

this phasing of the six element phased array "antenna" produces the typical two lobe far field radiation pattern. As a laser, each stripe has gain at several distinctive wave-lengths. Each wavelength is phase locked across the stripes by evanescent coupling between adjacent stripes. The individual stripe radiation angular divergence is about 20 degrees by 40 degrees. Running down the lists here to make sure that I've covered everything, the fiber was a Spectran 75, with 75 micron core and 125 micron cladding. It's a step index fiber therefore it is multimode and non-polarization preserving. Its numerical aperture for you optical people is 0.292. The diode is near field coupled, and the transmitting lens images the fiber output, projecting it on a space-like hypersurface at infinity. Translated that means the lens focuses the output surface of the fiber at infinity.

[ST CTF vgSpect] Here is a typical spectrum output of the laser diode showing the kind of wavelengths or frequencies that you're going to see for one of these diodes. As I said, this is a multispectral diode due to the fact that you can support several modes of oscillation inside a laser cavity of the dimensions used in this laser, because the laser cavity is fairly large.

[ST CTF vg9] Here's how we made the measurements. One of the differences between the setup shown and setup used to produce the data I will present is that the actual space qualified laser diode modules were not used. Measurements of the actual space qualified laser diode modules were made and were not significantly different from the data you will see. The laser bias supply drove the laser, which drove the optical fiber and transmitting lens. The Fourier lens is used to transform to the far field here at the detector. The far field is sampled through a pin hole whose 50 micron diameter approximates the receiver aperture at the normal operating range of the communication link. The pin hole and avalanche photodiode detector (APD) are mounted on a Klinger motorized scanning $x - y$ translation stage. The APD is a detector with internal gain. As such it had sufficient signal output to directly drive a 12 bit analog to digital converter. This was all controlled by a HP-9000 Series 220 computer. With this setup it was possible to automate the measurement of the far-field at the 50 micron aperture over a 120 point on a side square grid.

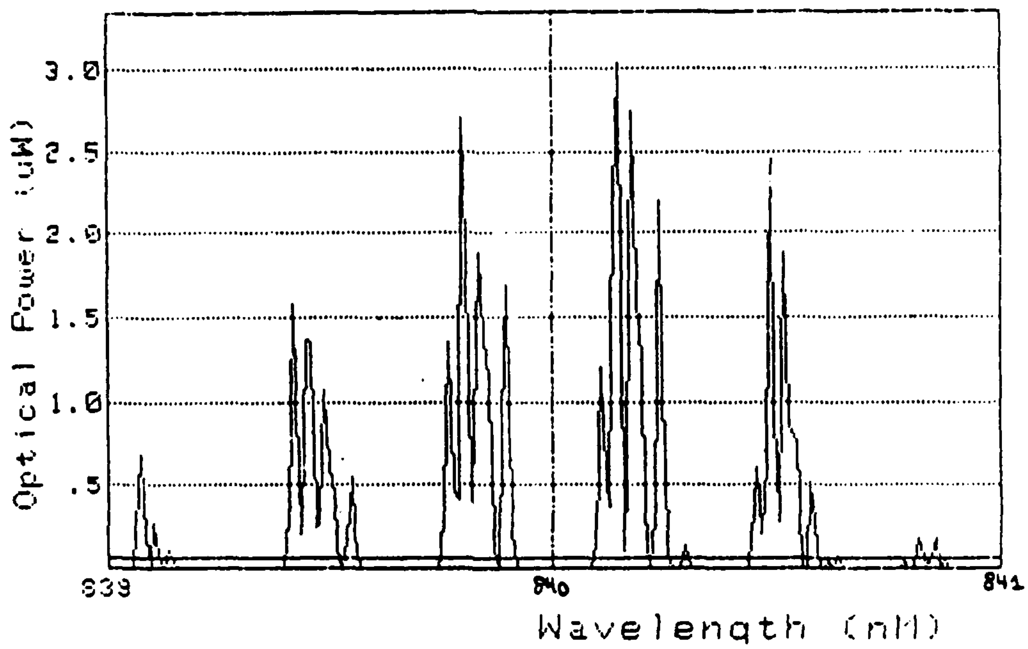
[ST CTF vg10] We measured that 120×120 grid in the far field, a measurement of a far field pattern that is essentially circular. From this circular intensity distribution (converted to power by sampling through a finite aperture) we took a 64×64 subarray from the center, indexed by integers m and n on 50 micron steps. From this I've computed the 2-D FFT directly, the probability density function, the 2-D autocorrelation, the 2-D autocovariance, the 2-D power spectral density from the autocovariance (rather than from the autocorrelation), and the space-time correlation lengths from the autocorrelation function.

[ST CTF vgPedestal] Now we can go fairly rapidly here and flip through 2-D plots that convey little information but that show that we did take the data on the basis that I said. Here's the 120×120 point grid. This spike in the back of the plot is a reference level, and the pedestal with the jagged top is a typical intensity plot.

[ST CTF vg64Subset] We take a 64×64 subset of that same data, and it looks like that when you plot it out with hidden line suppression.

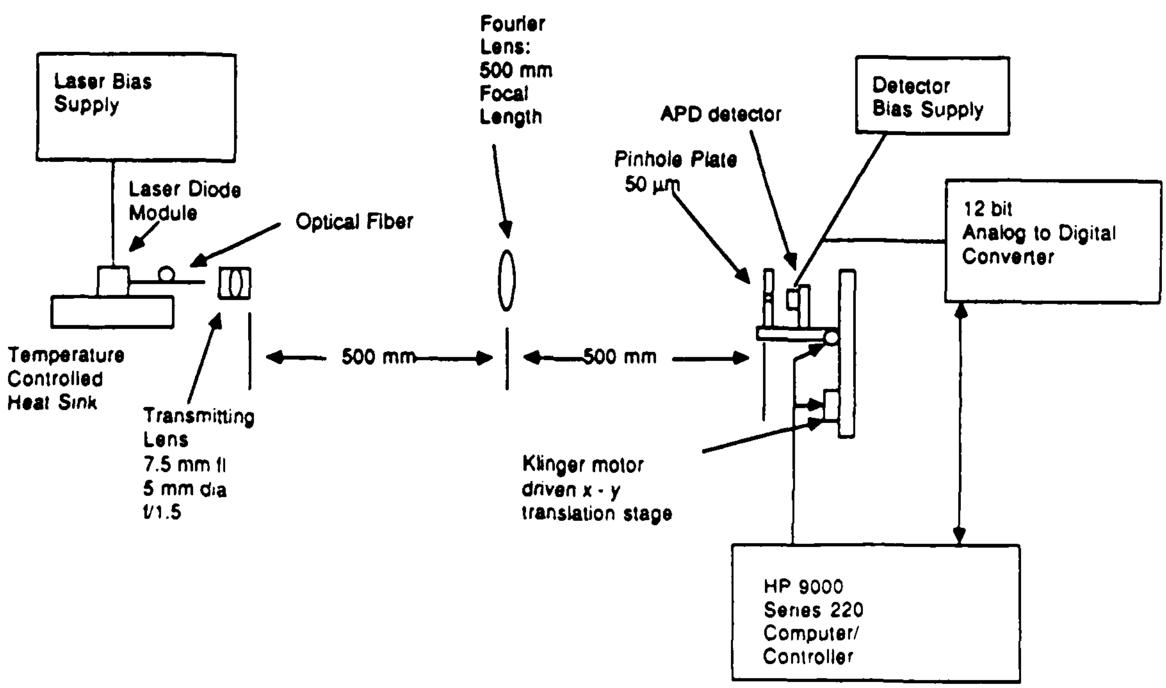
[ST CTF vgpdf] When you take the whole pedestal that you saw earlier and plot the

TYPICAL SPECTRUM OF LASER DIODE w/ OPTICAL FIBER



ST CTF vgSpect

SPECKLE TEST MEASUREMENT SETUP



ST CTF vg9

DATA TAKEN

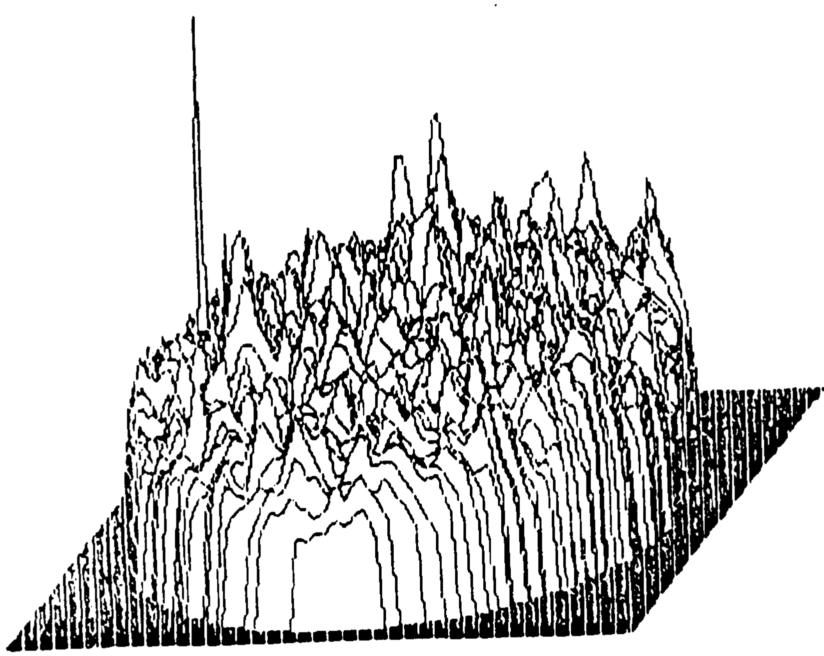
- MEASURED THE POWER THROUGH A 50µm APERTURE OVER A 120 x 120 GRID IN THE FAR FIELD
- HAVE A 64 x 64 ARRAY OF DATA TAKEN FROM THE CENTER OF THE 120 x 120grid:

  $p(m,n)\ :\ (m,n) \in (1\ldots64;\ 1\ldots64)$

- FROM THIS COMPUTED
  - -- 2-D FFT OF THE DATA DIRECTLY IN A 64 x 64 SUBSET
  - -- PROBABILITY DENSITY FUNCTION: MEAN, VARIANCE, SPECKLE CONTRAST
  - -- 2-D AUTOCORRELATION: 64 x 64 --> 32 x 32 SUBSET
  - -- 2-D AUTOCOVARIANCE: 64 x 64 --> 32 x 32 SUBSET
  - -- 2-D SPATIAL POWER SPECTRAL DENSITY FROM THE AUTOCOVARIANCE
  - -- THE ST CTF CORRELTATION LENGTHS FROM THE AUTOCORRELATION

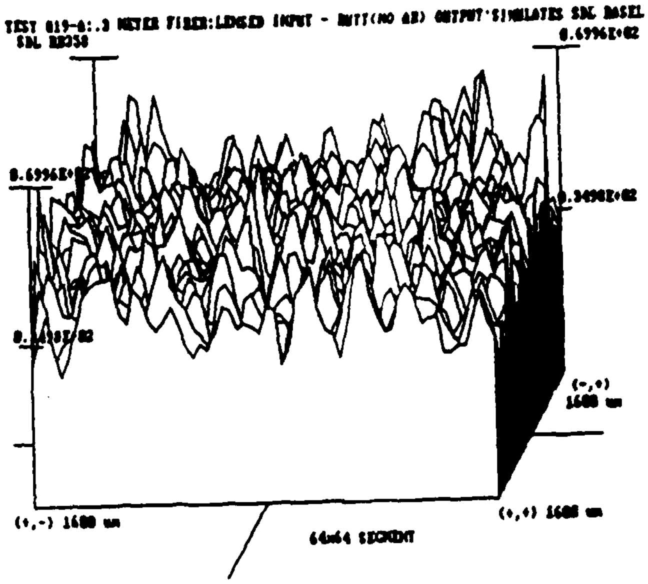
ST CTF vg10

TYPICAL INTENSITY PLOT OF RAW DATA IN 120 x 120 GRID



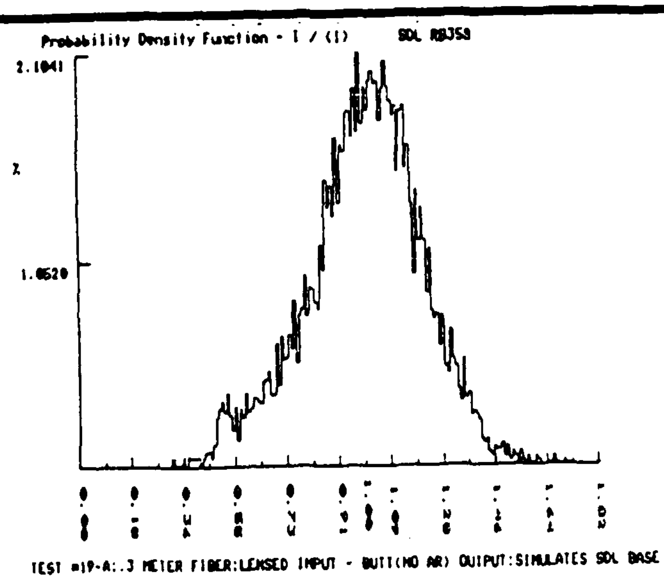
ST CTF vgPedestal

TYPICAL INTENSITY PLOT OF RAW DATA IN 64 x 64 GRID



ST CTF vg64Subset

PROBABILITY DENSITY FUNCTION



TEST #19-A:.3 METER FIBER:LENSED INPUT - BUTT(NO AR) OUTPUT:SIMULATES SDL BASEL

SDL RB358

The mean value = 1918.940975
The mean- squared value =3839351.500000
The sigma = 396.214996
The contrast = 0.206475

ST CTF vgpdf

probability density function, this is what you find. Here is the mean value at 1.0. The pdf is plotted around that value using histogramic type of plotting. Originally the pdf was not used. We only started considering the probability aspects of the problem when we began working on the communication performance prediction consequences of speckle. This is what a communication engineer needs for the speckle because what you usually end up computing is a bit error rate conditioned on a given received signal level. Then you have to treat that as the conditional computation it is, and use this probability density function on the intensity to compute the average bit error rate and probability of delivered service. The pdf was a new result when we computed it. When one has the pdf, one can compute the contrast, a statistical measure of the variability of the speckle pattern used by experts in that field. I will not go into this further.

[ST CTF vg11] Here is the notation for the FFT we took, the normal type of transform, expressed in DFT form.

[ST CTF vgRawFFT] Here is the plot of the magnitude of the 2-dimensional FFT. This was rearranged to plot zero frequency at the center of the plot, so the spikes at the center denote low spatial frequencies - where frequency is expressed in cycles per meter. You can see that there is significant power in the higher frequencies meaning that the 2-D surface will be "rough". Note that from an analysis point of view, that this type of plot is meaningless when attempting to determine statistical properties of a 2-dimensional random process.

[ST CTF vgACorr] The proper thing to be computing when analyzing statistical properties is the autocorrelation or autocovariance, and from that the power spectral density. When we take the autocorrelation of this data, you see a central spike which you would hope to see, but you can tell by the shape here that it's not well correlated with itself as you try to move away from a given point. This is a measure of how well distant values of intensity are correlated with the value you just left. So this is the autocorrelation function, ...

[ST CTF vgACov] which is almost indistinguishable from the autocovariance for this particular speckle pattern. That means that the subarray sums were approximately the same over the pattern. So this is what happened before we learned of Bill's model.

[ST CTF vgSpPSD] We computed the spatial power spectral density of the data and were able to get the smoothing Bill has referred to earlier. Thus we could estimate how much power is in the higher frequencies of that particular intensity fluctuation pattern.

[ST CTF vg12] This chart defines the exact processing that was used to compute the autocorrelation and the autocovariance. We computed the spatial power spectral density with the FFT using the formulation that I had described earlier.

[ST CTF vg13] Finally, we went ahead and learned Bill's model. He gave us a paper from MILCOM '88 and we went ahead and computed the correlation lengths L_x and L_y using this discretized form of the integrals. Again we used the normalized ρ, the normalized space-time correlation function, since that's where the information is about the pattern.

[ST CTF vg14] We computed the correlation lengths shown. We had actually analyzed two distinctly different speckle patterns, as the real reason for doing this work was to find ways to reduce the speckle. One was generated by the step index fiber system that was diagrammed in this talk, the other

DATA PROCESSING

- **HAVE A 64 x 64 ARRAY OF DATA**

  $p(m,n) \; : \; (m,n) \in (1 \ldots 64; \; 1 \ldots 64)$

  WHERE WE MEASURED VOLTAGES AT EACH POSITION. THE VOLTAGES ARE RELATED TO THE POWER BY

  $$v = I \, R_L = p \; Ro \; M \; R_L$$

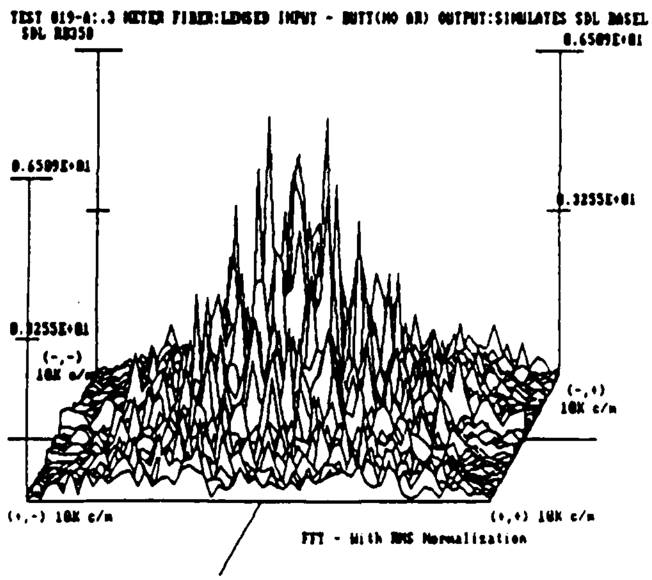
- **FFT (In DFT form)**

  $$\mathcal{P}(k,l) = \sum_{m=1}^{64} \sum_{n=1}^{64} p(m,n) \, e^{-j \, 2\pi \, (m-1)/64} \, e^{-j \, 2\pi \, (n-1)/64}$$

  THIS WAS REARRANGED TO PUT ZERO FREQUENCY AT THE CENTER OF THE PLOT.

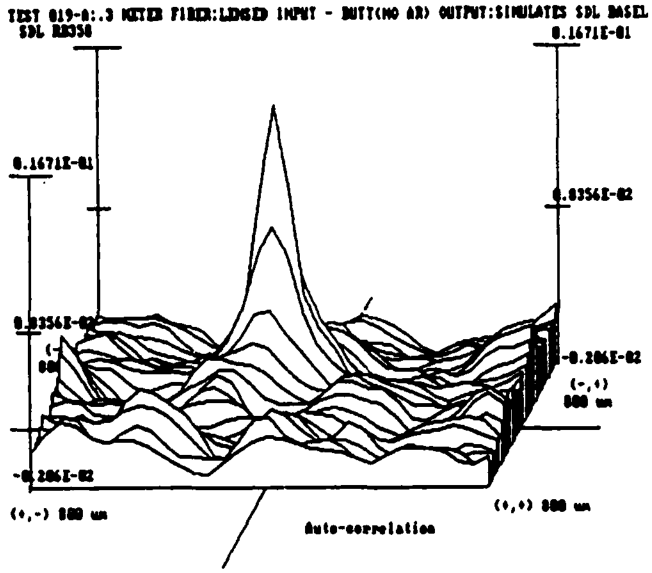
ST CTF vg11

FFT OF INTENSITY DATA



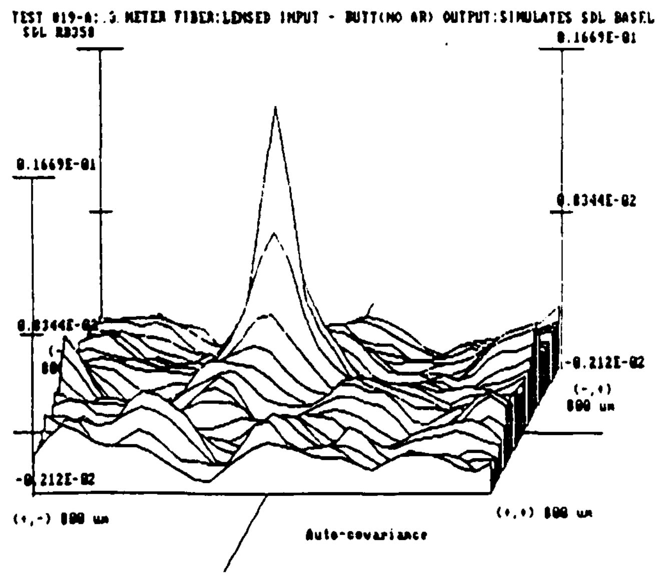
ST CTF vgRawFFT

AUTOCORRELATION OF DATA



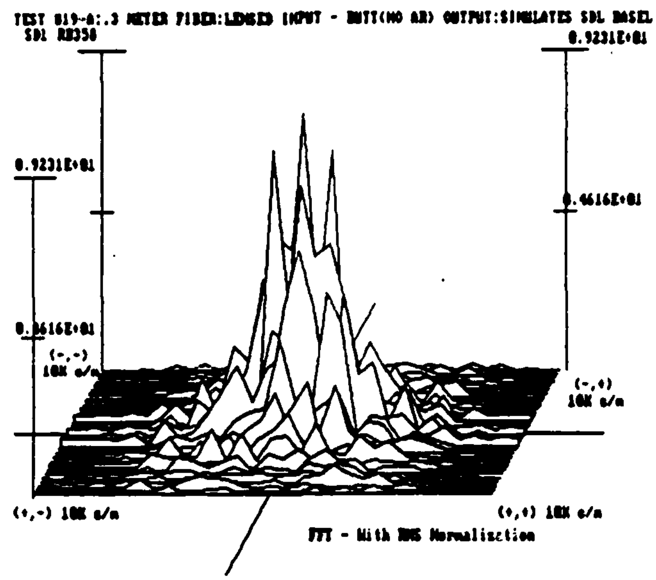
ST CTF vgACorr

AUTOCOVARIANCE OF DATA



ST CTF vgACov

SPATIAL POWER SPECTRAL DENSITY OF DATA



ST CTF vgSpPSD

DATA PROCESSING (Continued)

- **AUTO CORRELATION**

$$\mathfrak{R}(k,l) = \frac{1}{(32)^2} \sum_{m=1}^{32} \sum_{n=1}^{32} p(15+m,15+n)\, p(m+k,\, n+l) \qquad \text{PEAK AT } (15,15)$$

NORMALIZED AUTOCORRELATION:

$$\rho(k,l) = \frac{\mathfrak{R}(k,l)}{\mathfrak{R}(15,15)}$$

- **AUTOCOVARIANCE**

$$\mathfrak{S}(k,l) = \mathfrak{R}(k,l) - \frac{1}{(32)^2} \sum_{m=1}^{32} \sum_{n=1}^{32} p(15+m,15+n) \sum_{m=1}^{32} \sum_{n=1}^{32} p(m+k,\, n+l)$$

- **SPATIAL POWER SPECTRAL DENSITY**

2-D FFT OF $\mathfrak{S}(k,l)$

ST CTF vg12

ST CTF MODEL PARAMETERS

- **CORRELATION LENGTHS FROM THE AUTO CORRELATION**

$$L_x = \sum_{k=1}^{29} (k-15)\,\Delta x \rho\,(k,15)$$

$$L_y = \sum_{k=1}^{29} (k-15)\,\Delta y \rho\,(15,k)$$

ST CTF vg13

RESULTS OF COMPUTATION

- **CORRELATION LENGTHS**

	L_x	L_y	N_x	N_y
ISL	- 21.9 μm	- 174.45 μm	- 0.438	- 3.489
GRIN/LOOP	272.76 μm	146.79 μm	5.455	2.935

- **INTERPRETATION**

An N LESS THAN 1 INDICATES A CORRELATION LENGTH LESS THAN THE 50 μm STEP SIZE AND THE 50 μm APERTURE OF THE DETECTOR, WHICH I INTERPRET AS A DELTA FUNCTION

THE COMPUTATION OF CORRELATION LENGTH COULD OBVIOUSLY BE DONE FOR VALUES OF x AND y OTHER THAN THE CENTER VALUE (A SORT OF MARGINAL CORRELATION LENGTH).

IT IS NOT CLEAR THAT THE CORRELTATION LENGTH SHOULD NOT BE COMPUTED AS A RADIAL VALUE, GIVEN THE SYMMETRY OF THE SYSTEM. ARE WE MAKING SOME IMPLICIT ASSUMPTION ABOUT A FUNCTIONAL RELATIONSHIP BETWEEN CORRELATION LENGTHS MEASURED ALONG ORTHOGONAL DIRECTIONS? A RADIAL COMPUTATION WOULD ALLOW US TO PUT SUCH AN ASSUMPTION TO THE TEST (OR TO TEST THE DATA AGAINST SUCH AN ASSUMPTION).

ST CTF vg14

CONCLUSION

- GIVEN THE STATIC DATA WE HAVE TAKEN, A BEGINNING HAS BEEN MADE IN APPLYING THE ST CTF MODEL TO IT
- SOME DYNAMIC DATA HAS BEEN TAKEN AND REPORTED BY BREMMER IN MILCOM '88 PAPER 32.6
- A COMPLETE SET OF DYNAMIC DATA HAS NOT BEEN TAKEN OVER THE BEAM. AS THE SYSTEM IS ORKING, WE DO NOT ANTICIPATE TAKING SUCH DATA.
- THE ANALYTICAL EVALUATION OF THE IMPACT ON COMMUNICATION SYSTEM HAS YET TO BE DONE.

ST CTF vg15

was generated with a graded index (GRIN) fiber with a storage loop made by splicing a loop into the system using two fiber optic Y-couplers. (This is actually a very interesting system from a theoretical viewpoint because of the Markov channel-with-memory considerations which result from photonic storage in the optical fiber loop.) Because of time constraints I've suppressed showing you this second case. The system that was actually put into the intersatellite link has a correlation length of -21 micons in x and 174 microns in y. This corresponds to numbers of pixels, if you will, on 50 micron centers of 0.4 and 3.48. How does one interpret such a thing? Well certainly the correlation lengths can be negative in this case because we had a center at (0,0) in two dimensions. So you can have correlation lengths going either way. I interpret the N here as being less than 1 as indicating that you essentially have a delta-function. It's much less than the 50 micron step size and therefore I say I've got a delta function in that particualr direction.

Thus, if you look back, which you can't do because you don't have copies of the viewgraphs, [but now you do in this written form] everything was referenced to the center value x and y of the autocorrelation function. Obviously you can start choosing values away from the center and do a kind of a marginal computation of correlation lengths as you move away from the origin. Whether or not you ought to be actually doing a radial computation given the underlying cylindrical symmetry of the beam is a good question. These are the issues that are starting to arise as we start to try to apply this model.

[ST CTF vg15] So given that static data, we've made a beginning in applying the space-time correlation function model to it. As I mentioned, dynamic data was taken and reported by Bremmer in the MILCOM '88 paper. It would be good to have a complete set of such dynamic data. Unfortunately the program involved views their problem as being solved because they took the easy way out. They got the actual modem that drives the laser and demodulates the received signal, they set up a complete loop-back link and proved to themselves that with the gimbal jitters involved, and the waveform distortions involved, the system as designed delivered the performance that they required. So there's probably little motivation for taking the full set of what I'd call dynamic data over the complete 128 $\times$ 128 location intensity profile. We have not done any evaluation analytically of the impact on the communication system by using that probability density function or the outcome of the model. What the model will allow you to do is model the effect even in the static case of having gimbal jitter in this particular inter-satellite link. Because gimbal jitter is going to tend to make that field "dance" in the far field while your receiver is stationary with respect to that jitter, you'll get a time variation based, even in the static speckle case, on the gimbal jitter. Those things have not yet been computed, and we reserve that for future work. And that completes the presentation.

LINDSEY: The timing is just right. We are schedule to have our pictures taken and I don't know if the photographer is here or not. I know he's arrived outside but just how long it takes for him to set up the camera and all those good things, I'm not sure. I think it's going to be outside on the lake or something. So if you'd like to take a break and have coffee or whatever fruit is back there ... I think we should have our break, take the picture, and then come back and complete our discussions with regards to measurements and so on. It

looks like they took our goodies away. We abandoned them and they abandoned us so maybe we can get started yet.

O.K. we'll change the tempo, I think in terms of our discussion and now we'll talk about taking measurements with regards to characterization of various channel parameters and maybe whatever, and the first speaker after our break will be Paul Sass. Paul ....

PAUL SASS: *Wideband Channel Measurement Experience*

Thanks, Bill. As has been said a few times, we are going to change the tempo. I'm not going to compete with the theoreticians and put up very many equations. What I'm going to do is try to share with you some of the hard lessons we've learned over the last 5-10 years actually, in the field.

The material [SLIDES 1-2] I'm presenting here, admittedly, is past history. We have been doing this, as I said, over quite some time. But my motivation stems from a discussion I had with Mike Pursley about a year or two ago where he was lamenting the fact that we really didn't have the data to characterize the channel, whatever that means. I'm not going to answer it like Bill Lindsey, but what I'm going to do is to share with you some of the data we have acquired and put the onus on you to tell me how we can use this data to get these answers. I'm not going to propose getting a transmittance function specifically but you'll see what we did.

A little in the way of history. About 10 years ago, we at Fort Monmouth initiated a program to go out in the field and make wideband channel characterization measurements. Our interests were in the UHF band, 200 MHz to 2 GHz. We were interested in possibly or ultimately deploying spread spectrum systems in that frequency range. We had JTIDS coming along which was a frequency hopper over 250 MHz, centered around 1 GHz. We had a few other systems that were direct sequence, fairly nominal bandwidths and we were talking about packet radio at that time, using a direct sequence waveform of up to 100 MHz or so. So we needed wideband channel characterization data. In around 1982, we started a program which was competitively solicited and won by SRI International to design and build a wideband PN channel probe operating over that frequency band. At the same time we started with Al Schneider, who is joining me today, to look at one specific channel, the forested communication channel. He was involved from the beginning in designing the experiments and helping us analyze the data. Also, mid stream we got involved with Ray Luebbers through the Army Research Office who was interested in the Geometric Theory of Diffraction and how it might apply to wideband channels. He had only applied it to CW or narrowband signals up until that point. So he saw our measurement program as a unique opportunity to get some wideband data.

As I said, our work mainly stressed forest channels. The program in summary started with a prototype built by the people out at ITS, Boulder, Colorado, Dr. George Hufford and company. Bob Hubbard built a prototype in 1979-80 time frame which was a wideband PN probe and demonstrated its application for channel characterization. It was an analog system, very crude to use, but basically it proved that you could use a system like a PN probe to measure channels along the lines of Phil Bello's work much earlier than that. Over the period 1982-85, we actually built and tested and brought to the field an automated wideband measurement system to do that. [SLIDE 4] The system

WIDEBAND CHANNEL MEASUREMENTS *

-A TOUCH OF REALITY-

**IS IT REAL........OR IS IT MEMOREX?*

PAUL SASS
CENTER FOR C3 SYSTEMS
US ARMY CECOM
FORT MONMOUTH, NJ

SLIDE 1

PROGRAM HISTORY

OBJECTIVE-

CONDUCT A SERIES OF WIDEBAND CHANNEL CHARACTERIZATION EXPERIMENTS TO OBTAIN DATA SUITABLE FOR UHF SPREAD SPECTRUM SYSTEM PERFORMANCE ANALYSIS IN GROUND MOBILE ENVIRONMENTS.

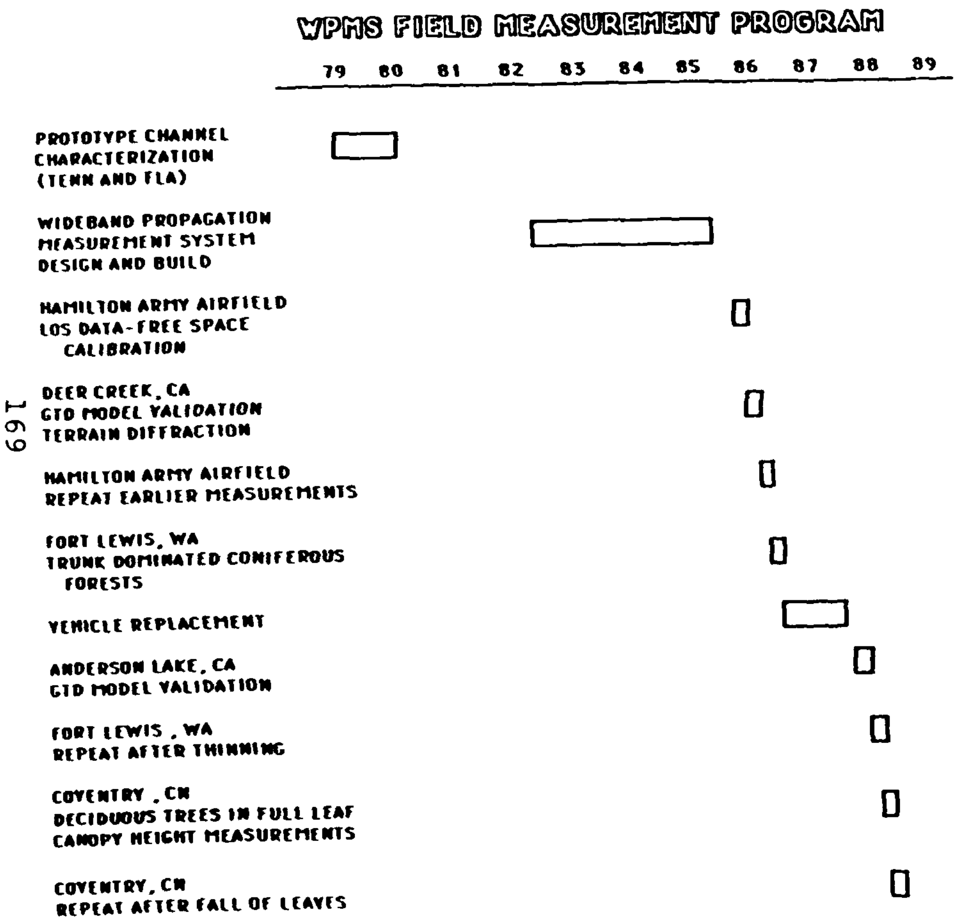
COMPONENTS-

WIDEBAND PROPAGATION MEASUREMENTS
- 1982-1988
- SRI INTERNATIONAL
- DESIGN, BUILD, AND FIELD WPMS
- DATA ANALYSIS

FOLIAGE RESEARCH/ MODELLING
- 1980-1989
- CYBERCOM CORP-DR. ALLAN SCHNEIDER
- BASIC RESEARCH IN FOREST MODEL DEVELOPMENT
- EXPERIMENT PLANNING
- DATA ANALYSIS

GTD MODEL VALIDATION/ APPLICATION
- 1986-1988
- UNIV OF PENN-DR. RAY LUEBBERS
- APPLICATION OF GTD MODEL TO WIDEBAND SYSTEMS

SLIDE 2



SLIDE 3

FIGURE 2 PHOTOGRAPHS OF TRANSMITTING AND RECEIVING VANS FROM CENTER OF 80

SLIDE 4

consisted of a transmitter in one vehicle and a receiver in another vehicle which I'll discuss in a few minutes. Over the next two years, following that point in time, except for a little break to deal with some vehicle problems that we had (and I'll talk about that), we compiled two years worth of a lot of data in a variety of sites. We concentrated on the forests in Fort Lewis, Washington which were largely trunk-dominated forests. This means the levels at which we were transmitting were largely through the trunks and not through the canopies. We also went to Connecticut for a different kind of tree, a deciduous tree, and examined the effects of leaves on propagating wideband signals.

As far as why we started, this is an old slide [SLIDE 5] I pulled out of my file from 10 years ago. At that time, as I said, we were talking about a generic spread spectrum UHF system, something called BIDS (battlefield information distributing system). It is not clear whether it would be a frequency hopper or direct sequence, but it would occupy significant bandwidth in the UHF band. And we asked all these questions: What happens when you put the systems in forests, for example? What happens when you put the signals through forests, what kind of multipath is measurable in the forest, what kind of delay spread is measurable in the forest and what other variabilities do you see in the forest? What are the parameters in the trees themselves that affect propagation?

The measurement approach itself relied on a wideband PN code PSK transmitted signal. At the receiver that signal is cross correlated with a reference, using what is called a sliding correlator which was based on Phil Bello's earlier work. The result is a time-varying impulse response of the channel. And if you keep track of things changing and how you do your sampling, you can get a fairly accurate representation of the time varying channel. This TVIR then becomes our measurement tool. It's an FFT away from getting the transfer function of the channel. What we measure is actually the output delay spread, I believe, based on Phil Bello's work. At the same time we use a lot of real time processing in the system to give the operator a real time sample of what's happening in the channel as well as setting ourselves up to do a lot of off-line data analysis. The system was heavily computer-controlled. The receiver system had an HP A-700 computer in it and enabled the operator to set up a whole experiment routine in advance and basically push a button causing towers to go up, antennas to rotate, frequencies to change, and things like that. It was very complicated, perhaps even too ambitious and you will see why in a few minutes.

[SLIDE 10] This is a block diagram showing the measurement approach itself. As I said, we generate a PN code, modulate an RF carrier, and transmit it over two transmitters simultaneously. We transmit on two channels and at the same time we receive over two parallel channels, each of which did a cross-correlation. Then we piece together the time-varying impulse response. Once the TVIR data was acquired, the data was dumped very quickly to a front-end memory in the computer and from there written to mag tape for subsequent processing. This slide is an example of what a TVIR looks like, obviously without any multipath. This is just a display of what a processed time-varying impulse response could look like and the corresponding FFT yields the spectrum of the received signal. Shown here is the full resolution of the system. I'll summarize the system capabilities in a minute.

WIDEBAND PROPAGATION MEASUREMENT PROGRAM

QUESTIONS BEING ASKED

- HOW FAR CAN BIDS CANDIDATE SYSTEMS COMMUNICATE IN FOLIAGE ?
- DO NARROWBAND PREDICTIONS APPLY TO SPREAD SPECTRUM ?
- WHAT BANDWIDTHS ARE SUPPORTABLE IN FOLIAGE ? (PULSE DISTORTION)
- WHAT IS UHF PROPAGATION MECHANISM IN FOLIAGE ?
- DOES RESOLVABLE MULTIPATH EXIST IN FOLIAGE ?
- WHAT IS COHERENCE BANDWIDTH IN FOREST ?
- WHAT ARE BEST ANTENNA HEIGHTS/POLARIZATIONS /PATTERNS
- WHAT ARE THE EFFECTS OF LEAF,GROUND,AIR MOISTURE, GROUND CONDUCTIVITY ?
- WHAT AFFECTS TIME VARIABILITY ?
- WHAT PHYSICAL FOREST PARAMETERS AFFECT FOREST COMMUNICATIONS/RANGING ACCURACY ?

SLIDE 5

WPMS MEASUREMENT APPROACH

- TRANSMITTED SIGNAL IS WIDEBAND PN CODE MODULATED PSK SIGNAL
- RECEIVED SIGNAL IS CROSS-CORRELATED AT RECEIVER WITH SLIDING REPLICA OF TRANSMITTED SIGNAL TO PRODUCE IMPULSE RESPONSE OF CHANNEL
- TVIR IS PRIMARY MEANS OF CHANNEL MEASUREMENT
- ALL DATA IS DIGITIZED, STORED IN REAL TIME IN FAST FRONT END MEMORY, THEN RECORDED ON MAGNETIC TAPE FOR SUBSEQUENT OFF-LINE DATA ANALYSIS
- REAL TIME PROCESSING PROVIDES IMMEDIATE OPERATOR FEEDBACK
- AUTOMATED CONTROL PERMITS EFFICIENT CONDUCT OF PRE-PLANNED EXPERIMENTAL PROCEDURES

SLIDE 6

The most general communication signal structure can be schematically represented by a time-versus-frequency diagram such as Figure 2. The instantaneous bandwidth, W, is hopped over some larger frequency

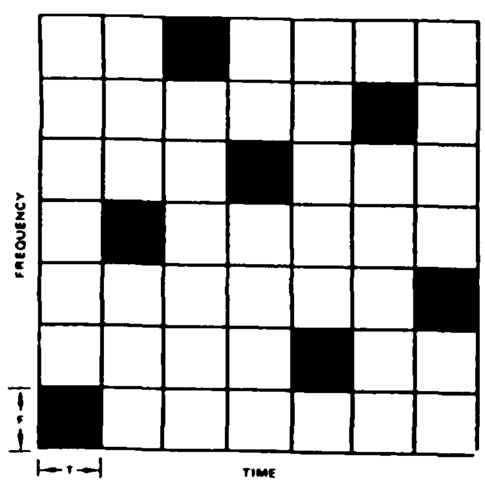


FIGURE 2 GENERAL COMMUNICATION SIGNAL STRUCTURE

space, and the interval, T, is the coherent processing interval. In general, TW >> 1. A channel can support this signal structure if and only if the chip interval, $\delta T_{d,} > {}^{1}/W$. Ideally, $\delta \tau_d > T$ and the frequency change $\delta f_d > W$, so that time and frequency hops are essentially uncorrelated. If the signals are appropriately combined (coded), improved performance results. On the other hand, as soon as $\delta \tau_d < T$ or $\delta f_d < W$, or both, performance rapidly degrades because the communication system cannot tell the difference between signal waveform (modulation) changes and channel-induced waveform structure.

SLIDE 7

E. Time-Varying Impulse Response

The intent of the system is to measure the time-varying impulse response (TVIR) or an equivalent functional representation such as the ODSF $[h(t,\xi)]$ of the "channel." Our WPMS design will measure a version of the ODSF. As discussed above, the ODSF and TVIR are entirely equivalent; the only difference is in the manner in which the time and delay variables are referenced. The delay variable, τ, of the ODSF is referenced (i.e., equals zero) to the time at which the impulse is applied to the communication channel. Hence, the ODSF is equivalent to an oscilloscope display of the response of the system to a unit impulse for which the time origin of the oscilloscope is set to the time of application of the impulse to the channel. This input time is the only place "absolute" time appears in the ODSF. The input-output relationship using the ODSF is

$$y(t) = \int h(t - \tau, \tau)\, x(t - \tau)\, d\tau \quad . \qquad (17)$$

The measurement concept is to transmit a continuous signal that is direct-sequence-spread over a 250-MHz characterization bandwidth (500 MHz null-to-null) by a pseudonoise (PN) sequence. The received signal is then given by Eq. (17). Cross-correlating the received signal with a replica of the original PN sequence gives the channel impulse response. That is, given a (baseband) PN probe signal, x(t), the transmitted signal has a complex representation

$$x(t) = \underline{x}(t)\, \exp(j2\pi f_c t) \qquad (18)$$

where f_c is the carrier frequency. The received signal is

$$y(t) = \underline{y}(t)\, \exp(j2\pi f_c t) \quad , \qquad (19)$$

and the ODSF, represented as a frequency-shifted baseband signal, is

$$h(t,\tau) = \underline{h}(t,\tau)\, \exp(j2\pi f_c t)) \qquad (20)$$

SLIDE 8

The received signal can be written as

$$y(t) = \exp(j2\pi f_c t) \int \underline{h}(t - \tau, \tau)\, \underline{x}(t - \tau)\, d\tau \quad . \tag{21}$$

Cross correlating with a shifted conjugate replica of $x(t)$ we have

$$\langle y(t)\, x^*(t - \xi)\rangle =$$
$$\exp(j2\pi f_c \xi) \int \underline{h}(t - \tau, \tau)\, \langle x(t - \tau)\, x^*(t - \xi)\rangle\, d\tau \quad . \tag{22}$$

For $\langle \underline{x}(t_1)\, \underline{x}^*(t_2)\rangle = \delta(t_1 - t_2)$, we obtain:

$$\langle y(t)\, x^*(t - \xi)\rangle = \exp(j2\pi f_c \xi)\, \underline{h}(t - \xi, \xi) \quad . \tag{23}$$

If the reference signal is phase locked to the received carrier [and does not shift as we shift $\underline{x}(t)$], the exponential term vanishes, and we are left with $\underline{h}(t-\xi,\xi)$.

The step from Eq. (22) to Eq. (23) implicitly assumes that the channel does not change during the measurement interval. In the simplest case, this is the time required to correlate a single PN sequence with itself. If several PN sequences are required to complete a measurement of $\underline{h}(t,\tau)$, then the channel must remain constant for this interval of time.

The variable, t, in all of the above expressions is defined as the current value of the time variable; hence, t is a running variable. If t_o denotes the time the unit impulse is applied to the system, then $t = t_o + \xi$ and $\underline{h}(t - \xi,\xi) = \underline{h}(t_o,\xi)$, which is the ODSF.

The function we measure will actually be a time-shifted version of the ODSF, which is offset in time by an unknown portion of the propagation delay because of the unsynchronized nature of the measurement system we have designed. We do not believe that measurement of propagation delay warrants the expense of a fully synchronized system.

SLIDE 9

The block diagram of Figure 3 shows the basic mathematical operations performed by our measurement system. The integration in the receiver requires one complete period of the PN sequence to obtain the

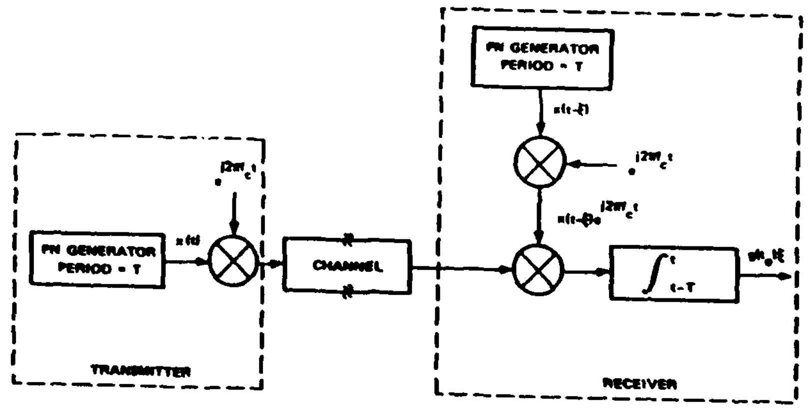


FIGURE 3 BLOCK DIAGRAM OF MEASUREMENT CONCEPT

value of $h(t_o,\xi)$ for a single value of the delay variable. Hence, if N values of ξ are required to adequately specify $h(t_o,\xi)$, then N periods of the PN sequence, or NT seconds, are required for the measurement of one complete sample of $h(t_o,\xi)$. For the measurement to be meaningful, $h(t_o,\xi)$ must not change over this time interval. The time, t_o, can be any convenient time during this interval; we shall define t_o to be the start of the measurement of a complete impulse response function.

SLIDE 10

In the forest channel, typical delay spreads looked like this. [SLIDE 11] The top one shows a delay spread of almost one and a half microseconds (1400 nanoseconds) with a vertical polarization antenna. With horizontal polarization antennas we typically saw much lower delay spreads. But in this case we see 200 nanoseconds of delay spread. We found very clearly that the forest did represent a complex scattering channel to the receiver.

As far as the system capabilities itself, we had a transmitter and a receiver, each mounted in a vehicle, integral self-elevating antennas that could raise each antenna up to 65'. We used two different antennas, omnidirectional and a 7 dB gain, log-periodic array covering the band 200 MHz - 2 GHz. We had 4 possible operating probe bandwidths ranging from unmodulated CW signals all the way up to 250 MHz direct-sequence spread spectrum. We also had a variety of code lenghts which enabled us to look at different multipath delay windows. And as I said, we were capable of transmitting and receiving on two simultaneous channels. The final system design resulting from an agonizing process over the years was really a compromise between our desire to achieve a total path-loss capability of 155 dB and yet achieve both the resolution and delay-spread capability over the full characterization bandwidth. We achieved all the desired parameter ranges, but we were unable to simultaneously achieve full path-loss, full delay spread, and maximum resolution.

The approach we took in bringing the system out into the field, as I said, was fairly ambitious, and started with well-understood channels. Our first attempts to do system calibration and make sure that everything was working the way it should, were obviously over channels that were simple to measure: free-space and simple two path reflections off runways. We progressed gradually to more complicated channels like the forest. This occured, as I said, over a period of two years of separate measurements, all of which are documented to varying degrees in a whole number of reports. My interest in coming here was to offer these reports or any of the data that any of you might be interested in.

That was all the theory. In theory it should all work. When we took it into the field we had a lot of problems. I wanted to briefly review some of the experiences we had. I thought it would help some of the theoreticians here understand the difficulty in getting the measurement data to support their theoretical work.

I have some slides on a few of the problems, but let me just talk about the ones I don't have any slides on. Electromagnetic interference presented a far greater problem than we expected. As I said, we were trying to characterize hundreds of megahertz of bandwidth in the UHF band. When we took the system into the field, we had a lot of trouble with others in band emitters. We originally had the concept that Phil mentioned, of being able to excise small numbers of interference sources, but we found out that the interference environment was so severe that we had no chance of excising in band interference with tunable notch filters. As a result, most of the useful data we got was confined to the forest, where the sources of interference were basically shielded from the measurement.

Another problem is caused by the fact that the forests were extremely inhomogeneous. In fact, Al Schneider has spent a lot of time trying to characterize these forests in some measurable way. We've seen the same problem trying to make path-loss measurements in the real world. We've set up UHF networks of 40

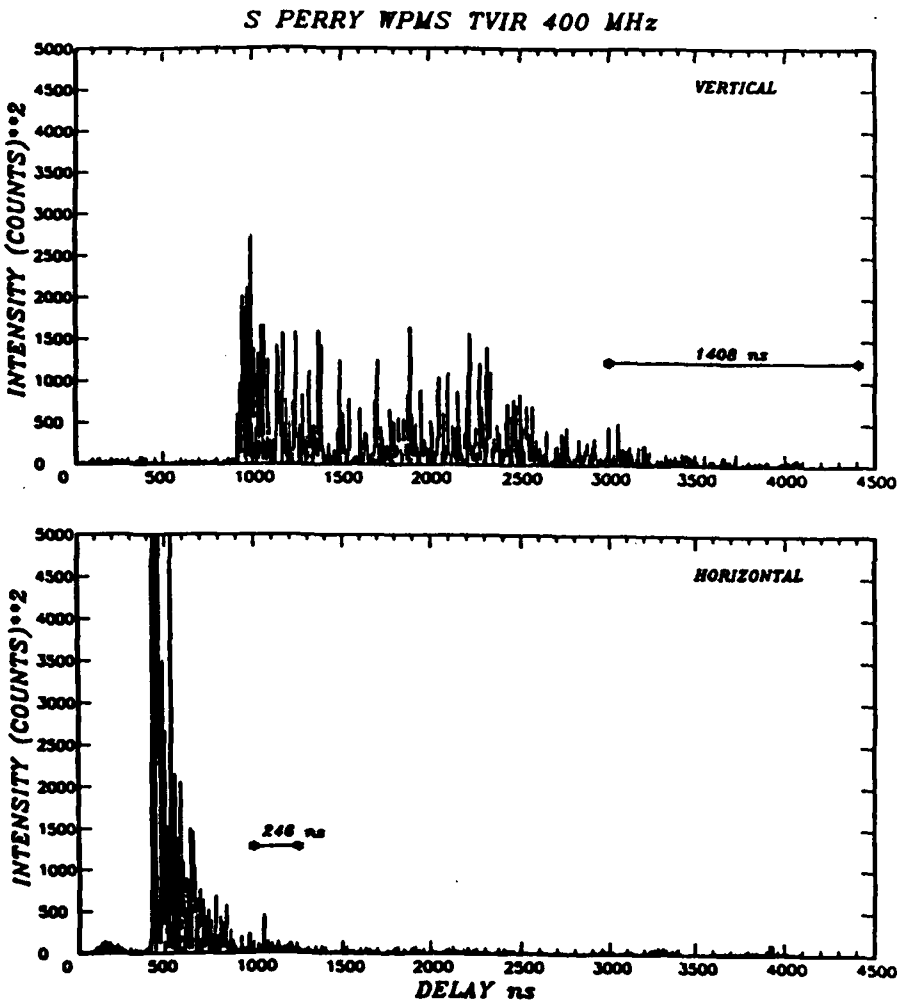


Figure 3.8 South Perry Woods Power TVIRs Measured at 400 MHz with Characterization Level and Calculated Delay Spread Indicated

SLIDE 11

WPMS CAPABILITIES

- MOBILE TRANSMITTED AND RECEIVER VEHICLES
- 65 FOOT SELF-ERECTING ANTENNA TOWERS INTEGRAL TO EACH VEHICLE
- TWO ANTENNA SYSTEMS--LPA AND BICONE
- OPERATIONAL FREQUENCY RANGE 200 MHZ TO 2 GHZ
- FOUR PROBE SIGNAL OPERATING MODES-CW, 50MHZ, 125 MHZ, AND 250 MHZ PN CLOCK
- FOUR PN CODE LENGTHS-255, 511, 1023, 2047 CHIPS
- TRANSMIT AND RECEIVE TWO SIMULTANEOUS PROBE SIGNALS ON TWO CHANNEL XMIT/ RCVR
- 100 WATT TRANSMITTER DESIGNED FOR FULL MEASUREMENT RESOLUTION OVER 155 DB PATH LOSS

SLIDE 12

or 50 radio nodes attempting to compare predictions and path-loss measurements over the links in the 40-node network, and we have had a lot of problem comparing the theory to actual practice. The path-loss prediction models that we have to deal with today are semi-empirical in nature and produce probabilistic predictions. Designing a statistically valid mesurement to compare to a sampled probability distribution proved difficult.

[SLIDE 13] "Mobile" means you need vehicles. We had a lot of troubles with vehicles alone. I'll show you how the vehicles evolved. As I said, we started out with a concept 10 years ago of small vehicles that can get in and out of forests and move in and around terrain. You'll see what we ended up with.

[SLIDE 15] Cost was another problem. As far as cost, I thought I'd just show you the price tags on some portions of the measurement program itself. This doesn't even include the data analysis that we've done and all the work that Al Schneider has been doing with the data for the forested channels that he's interested in. But as you can see, by the time we bought the measurement system, took care of the vehicular problems and ventured into the field, we were left with a fairly expensive proposition. This is particularly frustrating since it didn't leave us enough budget to really analyze the data as thoroughly as we'd like.

[SLIDES 16 ..19] This is a photo of part of our data, currently in storage at SRI International. We had a lot of philosophical arguments over the years about the need to preserve raw data as opposed to reduced data. We've done both. We've not only compiled data summaries of the raw data, but we've also preserved the raw data itself. We've estimated that 7000 Gigabytes of raw data were acquired in this measurement phase in a two and a half year period. In Fort Lewis alone we wrote 50 mag tapes worth, or 15 gigabytes of digitized data. As the system compiled its TVIR, it wrote the results into a 2 MB front-end memory and subsequently dumped it to tape for analysis. The raw data acquisition took anywhere from 4 seconds at full bandwidth spread, to about 4 minutes of CW data. So that was the range of data we could acquire before we'd fill our front end and have to dump it to tape. It wasn't quite as dynamic and as easily handled as we had thought when we first started.

I wanted to review the data processing burden itself. Maybe you have a feel for the data analysis problem. The raw TVIR data, as I said, was first written to tape. The first process in the data analysis routine was to correct for imperfections in the hardware. The way the correlation receiver is structured, it uses 4 separate correlator channels, each of which have I and Q channels. The phase errors between those correlator channels result in errors in the resulting TVIR. An off-line analysis routine was written to remove those phase-offsets on each of the 8 channels. This produces a corrected complex TVIR. That was the first objective of the measurement itself.

Once we had obtained the corrected TVIR, we did a lot of summary data file manipulations. We computed the average power TVIR and the spectrum, and wrote that to a summary data file which included all the experiment descriptive parameters, tree parameters, etc., everything that was stored in the experiment set-up. That makes up the part of the files that we ship around to most of the people who are looking at the data.

The raw mag tapes were processed off-line at SRI and then written to a summary data file on an HP mag tape. In some cases they

ENTER REALITY

COST OF FIELD EXPERIMENTS

DATA RECORDING/ PROCESSING ANALYSIS/ ARCHIVING

EQUIPMENT RELIABILITY

ELECTROMAGNETIC INTERFERENCE

MEASUREMENT OF INHOMOGENEOUS, NONDETERMINISTIC PHENOMENA

MOBILE MEANS YOU NEED VEHICLES

SYSTEM CALIBRATION/ NON-IDEAL MEASUREMENT TOOLS

HARDWARE CHALLENGES

SLIDE 13

EXPERIMENTS PERFORMED

FREE SPACE CALIBRATIONS

SIMPLE REFLECTIONS OFF SMOOTH RUNWAY

CW SPATIAL CORRELATION

SPREAD SPECTRUM SPATIAL CORRELATION

RECIPROCITY

FOREST LEAKAGE

STEM COUNT HYPOTHESIS

TRUNK DOMINATED FOREST BEFORE/ AFTER THINNING

TREE PROXIMITY

LATERAL WAVE

CROSS POLARIZATION

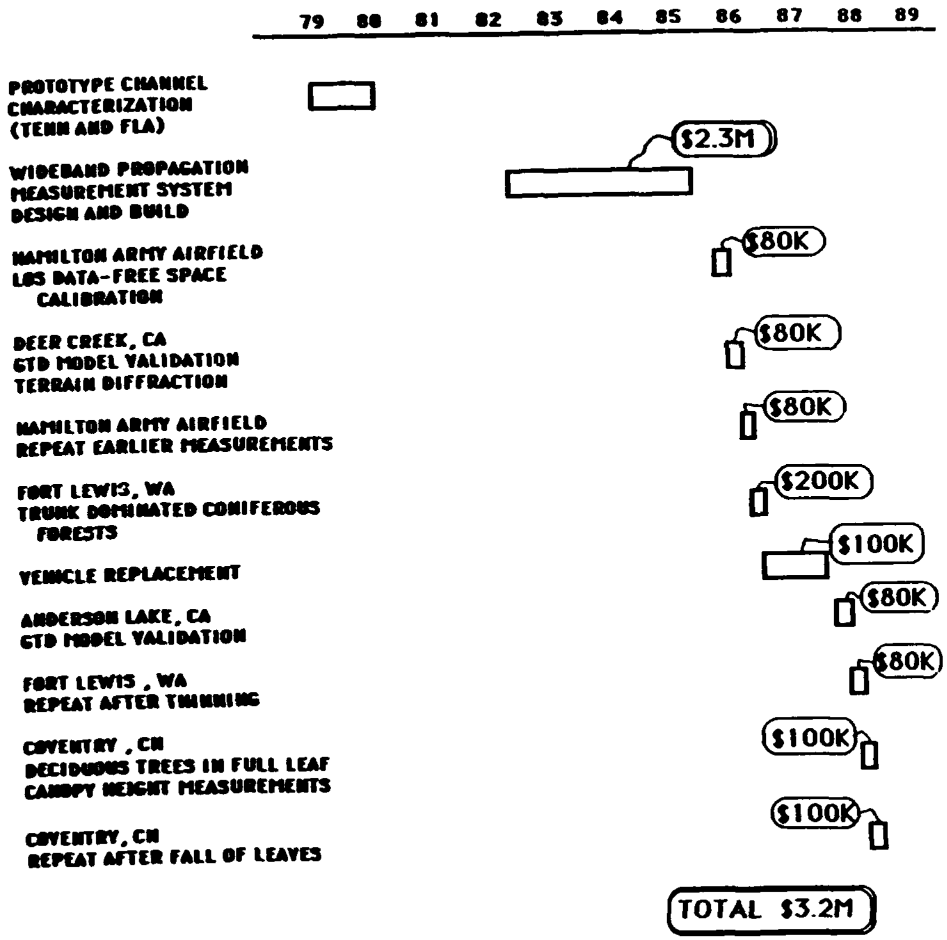
SIMILAR PATHS

CANOPY MEASUREMENTS IN DECIDUOUS TREES WITH/ WITHOUT LEAVES

SLIDE 14

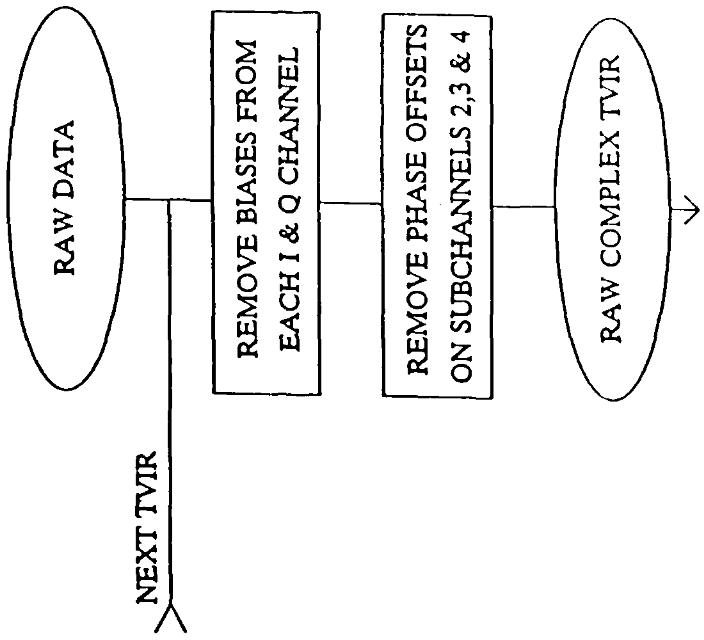
COST OF

WPMS FIELD MEASUREMENT PROGRAM



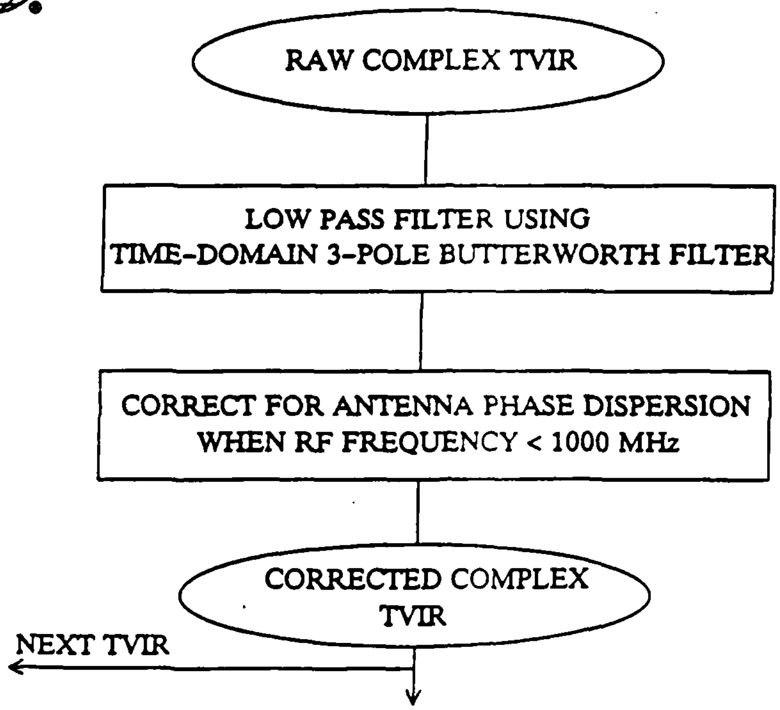
SLIDE 15

DATA PROCESSING SUMMARY



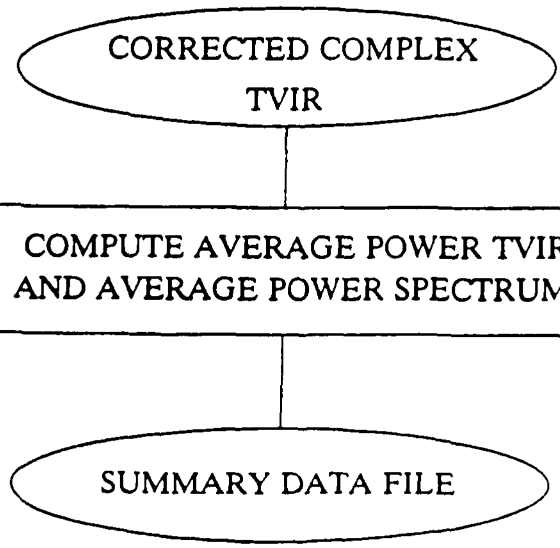
SLIDE 16

DATA PROCESSING SUMMARY (CONTINUED)



SLIDE 17

DATA PROCESSING SUMMARY (CONCLUDED)



SLIDE 18

WPMS DATA DISSEMINATION

WPMS SUMMARY DATA FILE
PRIME COMPUTER DISC FILE (BINARY)

SUMMARY DATA FILE
HP FORMAT, 9-TRACK TAPE (BINARY)

CECOM
ITS

ASCII SUMMARY DATA FILE
(PART OF HEADER, 1 OR MORE COMPLEX TVIRS
MINIMUM, AVERAGE, MAXIMUM POWER TVIRS
MINIMUM, AVERAGE, MAXIMUM POWER SPECTRUM)
PRIME COMPUTER DISC FILE (ASCII)

PENNSYLVANIA
STATE UNIVERSITY

ASCII SUMMARY DATA FILE
9-TRACT TAPE (ASCII)

ASCII SUMMARY DATA FILE
HP A700 COMPUTER DISC FILE (ASCII)

CYBERCOM

ASCII SUMMARY DATA FILE
HP CARTRIDGE (STREAMER) TAPE (ASCII)

SLIDE 19

FUN THINGS I DID WITH MY SUMMER*

GREG DESBRISAY
SRI INTERNATIONAL

- TAPE CONTROLLER FAILURES
- PN CODE GENERATOR FAILS—BAD SOLDER JOINT
- AC GENERATOR QUITS DUE TO UNVENTED FUEL TANK
- A/C HOT GAS BYPASS VALVE FAILS
- SYSTEM RELAY TELL-TALE LIGHTS INOPERATIVE DUE TO HIGH VEHICLE TEMPERATURES
- FAILED SYNTHESIZER IN XMTR—LOSE ONE CHANNEL
- PLL IN CALIBRATOR REPEATEDLY LOSES LOCK
- RECEIVE TOWER COLLAPSES THROUGH "DOWN" LIMIT SWITCH BENDING TOWER AND LPA ANTENNA
- KEYS LOCKED IN RENTAL VEHICLE
- CAR BATTERY DEAD ON RENTAL VEHICLE
- TWT FAILED—LOSS OF HIGH BAND OPERATION
- TOWER WINCH FAILS TO RAISE TOWER—HOMEBREW REPAIR
- OVERHEATED AC GENERATOR—KEEP DOORS OPEN
- CPU REFUSES TO BOOT AFTER GENERATOR QUIT
- LPA ELEMENT VIBRATES OFF—SPARE ANTENNA
- REAL TIME SOFTWARE MODIFICATIONS
- VEHICLE STUCK IN MUD
- REAL TIME I/O DISPLAY DIES—LOSS OF ONE CHANNEL REAL TIME OUTPUT

* 3 LONG DAYS AT FT. LEWIS, WA

SLIDE 20

were converted to ASCII summary files on various other media. At Penn State University Ray Luebbers worked directly from the ASCII summary file, but Al Schneider used an HP cartridge on a small HP 9816 computer. So we converted to a few different formats to provide it to as many people as possible.

I said something about the vehicles. This was originally the general concept of the transmitter vehicle. This was actually the vehicle we used back in the 1979-80 time frame, to hold the probe transmitter. You can see the omni-directional antenna on the roof. We actually went out and demonstrated that this system could work. The receiver at that time was a rented motor home that the folks at ITS had provided us. As shown here, one of the first lessons we learned was that you don't use guy wires in the woods. We had a lot of trouble with the towers in forests, trying to elevate antennas to the various heights. The tower we used 10 years ago needed guy wires and it was totally impossible to deal with. As a result our next phase requirements included measurement vehicles with integral, non-guyed towers. For a variety of reasons this is the ultimate transmitter vehicle we ended up with. As I said, it was a lot bigger than the GMC suburban we started with. After a number of attempts to design and purchase our own vehicles, we ended up with the vehicle provided at no cost from another contract. We modified it and installed the self-erecting towers that you see on the back and went to the field in Fort Lewis.

The receiver vehicle also got correspondingly bigger. I didn't even bring a picture of the receiver system. The receiver system includes several racks of HP equipment. It's fairly big and with all the creature comforts and the air-conditioning and all the support things that we needed to go into the field, the vehicle quickly grew. It was much bigger than we wanted, and it limits what you can do with a system like this, but we had to find ways of fielding this system. The reason we initially accepted such a big vehicle was that it was provided to us free by an Air Force Contract.

I thought you would appreciate seeing some of the problems we had in the field. You can read it yourself. This was put together at my request by one of the SRI engineers that ran the measurement phase. The real scary thing, if you look at the bottom, is that this represents a 3-day period in Fort Lewis, Washington.

[SLIDE 21] One of the other things I wanted to talk about was the problem caused by non-ideal measurement tools. There were two corrections that SRI had to implement in the data processing routines. The first was the TVIR calibration because of phase-errors between the correlator channels. They had to correct for that to come up with the corrected TVIR. The other thing is the antenna dispersion. The log-periodic arrays had non-linear phase over the significant bandwidth that we were measuring. Therefore, in order to get a corrected TVIR, we had to do some off-line calibration. In fact, SRI did a lot of work trying to correct for the actual antenna dispersion measured with the log-periodic array. [SLIDE 26] The result is shown by the dotted line of the correction. The solid line is the TVIR before the correction was applied. The dotted line shows a nice narrow correlation peak after correction. SRI found a fairly linear phase over the main beam of the antenna. Therefore, as long as you stayed within 45 degrees of the main beam of the antenna, the correction was valid. They also found that the correction was more necessary at lower

NON-IDEAL MEASUREMENT TOOLS--

CORRECTIONS FOR
- TVIR CALIBRATION-
-ANTENNA DISPERSION-

SLIDE 21

I N H O M O GE NE OUS

R A N D OM

M E D I A

SLIDE 22

COMPLEX TVIR CALIBRATION

To verify operation in the narrowband mode, the transmitter and receiver synthesizers are phase locked. A CW signal offset by δ Hz then gives the following outputs:

$$I_\ell(n;t_e) = A_\ell \cos\left[2\pi\delta n(mT+\Delta t)+\phi+\phi_\ell\right] + B_\ell \qquad (2.1a)$$

and

$$Q_\ell(n;t_e) = A_\ell' \sin\left[2\pi\delta n(mT+\Delta t)+\phi+\phi_\ell'\right] + B_\ell' \quad . \qquad (2.1b)$$

Here we have extracted the amplitude and phase of the complex gain, $E_\ell = A_\ell \exp\{i\phi_\ell\}$. A straightforward curve-fitting procedure is used to measure A_ℓ, δ, and B_k. Correlation of I_ℓ and Q_ℓ with sine waves of frequency δ are then used to determine ϕ_ℓ and ϕ_ℓ'. Proper operation of the correlator quadrature mixer requires ϕ_ℓ and ϕ_ℓ' to agree to better than 5°, which is the manufacturer's specification for the devices. Manual adjustments to the correlator are made to equalize the channel amplitudes to within a few percent and to minimize the bias levels. Any residual bias or amplitude imbalances must be corrected in the post-episode data analysis.

To calibrate the wideband operation of the system, the transmitter and receiver are operated in their normal operational modes but connected directly by cable to bypass the antenna systems. If the path length changes at a rate v, the received signal will be Doppler shifted by

$$\delta f = f_c \, v/c \quad \text{Hz} \qquad (2.2)$$

SLIDE 24

51

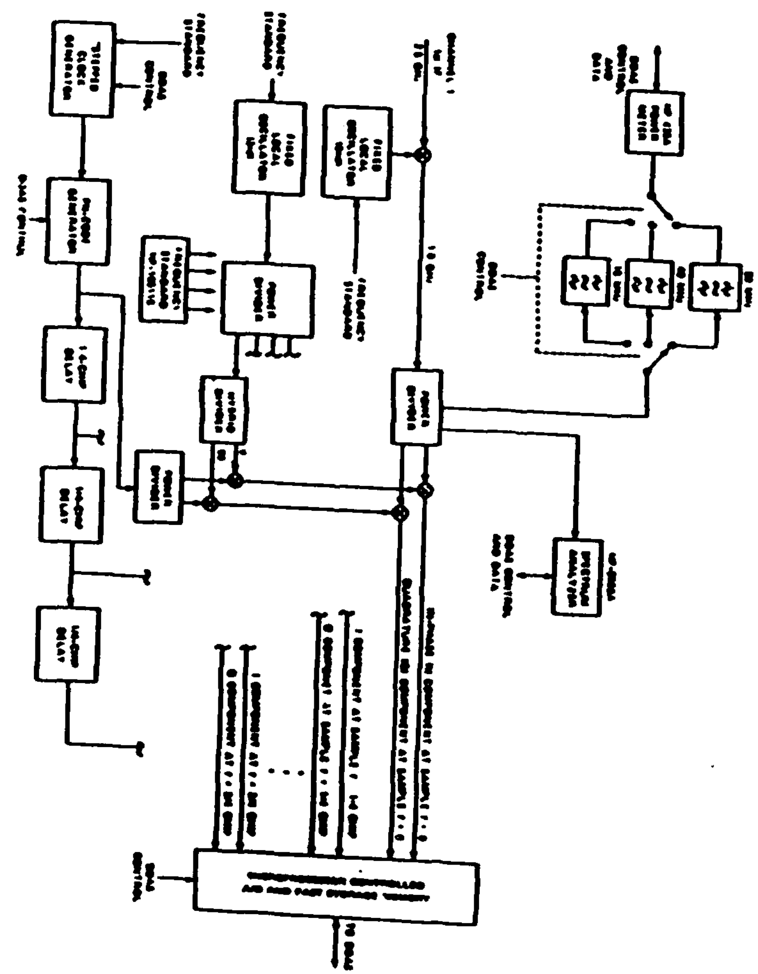


FIGURE 10 CORRELATOR, BLOCK DIAGRAM (SINGLE CHANNEL)

SLIDE 23

with the corresponding delay change

$$\frac{d\tau}{dt} = v/c = \delta f/f_C \qquad (2.3)$$

where f_C is the carrier frequency of the signal. We can simulate such a signal by offsetting both the rf and the code oscillators appropriately. For example, if the carrier frequency is 1000 MHz and it is desired to simulate a 4 Hz Doppler shift, the PRN generator's 250-MHz reference must be offset by (4Hz)(1000 MHz)/(250 MHz) = 16 Hz. It is not necessary to simulate the actual Doppler shift, however, and the RF and code offsets can be set to any setting that is convenient for the bias estimates.

The effect is to introduce a sinusoidal amplitude variation and a slow delay change into the complex signal. Formally, $R(t_n)$ in Eq. (2.2) is replaced by

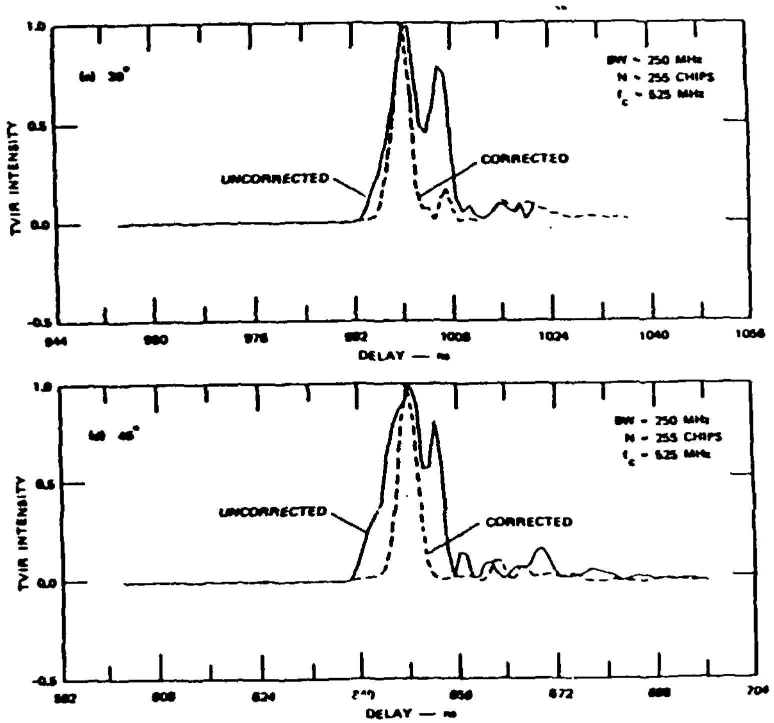
$$R(t_n - \tau(t)) \exp[2\pi i \delta f t] \quad . \qquad (2.4)$$

If we record the I and Q outputs at a constant delay, we can measure the complex envelope of the received signal in each receiver subchannel. This allows us to check the consistency of our CW measurements as well as measure the relative delays between the four subchannels.

To automate this procedure, we compute the cross-correlations between the subchannels to determine the subchannel delays. The phase offsets are calculated by correlating each I and Q signal with a sine wave of frequency δ(rf offset), as in the CW measurements. These phases are then used to determine the phase differences between the I and Q signals and between the subchannels. In principle, we could compensate for subchannel delay errors in postprocessing, but that would involve a time consuming interpolation. Thus, the results of our preliminary measurements were used to adjust the delays to within a small fraction of their nominal quarterchip values by trimming the lengths of the

SLIDE 25

CORRECTION FOR ANTENNA DISPERSION



SLIDE 26

4 DISCUSSION AND CONCLUSIONS

The WPMS system can resolve signal-delay structure on the order of 3 to 4 ns. Thus, very small changes in the geometry of the measurement or changes in the channel scattering characteristics can have a very large effect on the measurement. It is, therefore, important that all systems effects that can mask such changes be identified and, to the extent possible, compensated. In a companion report*, we evaluated a calibration procedure to remove receiver-induced biases. Effectively, the system can measure complex TVIRs with resolution approaching the ideal TVIRs filtered to remove spectral content beyond the second null in the ideal sinc^2 power spectrum.

The high gain antenna system, however, is highly dispersive and without compensation would severely limit the useful measurement capability of the WPMS system. Thus, our most important result in this report was to develop a dispersion correction scheme. We showed that a single-phase dispersion function can be used to recover the achievable resolution fully of the system for all antenna combinations. Thus, the delay resolving capability of the WPMS system can be fully exploited.

In the complicated scattering channels of interest, however, scattering from wide angles will be important. Under such conditions, the directional gain patterns of the antenna can affect the measured TVIRs significantly. We have provided a mathematical framework within which these effects can be evaluated, and provided an estimate of the main lobe of the antenna gain function for use when explicit computations are required. We expect the results to become more important as the planned measurement series progresses.

---

*Op. cit.

SLIDE 27

HARDWARE CHALLENGES

A VARIABLE CLOCK RATE, VARIABLE LENGTH CORRELATOR

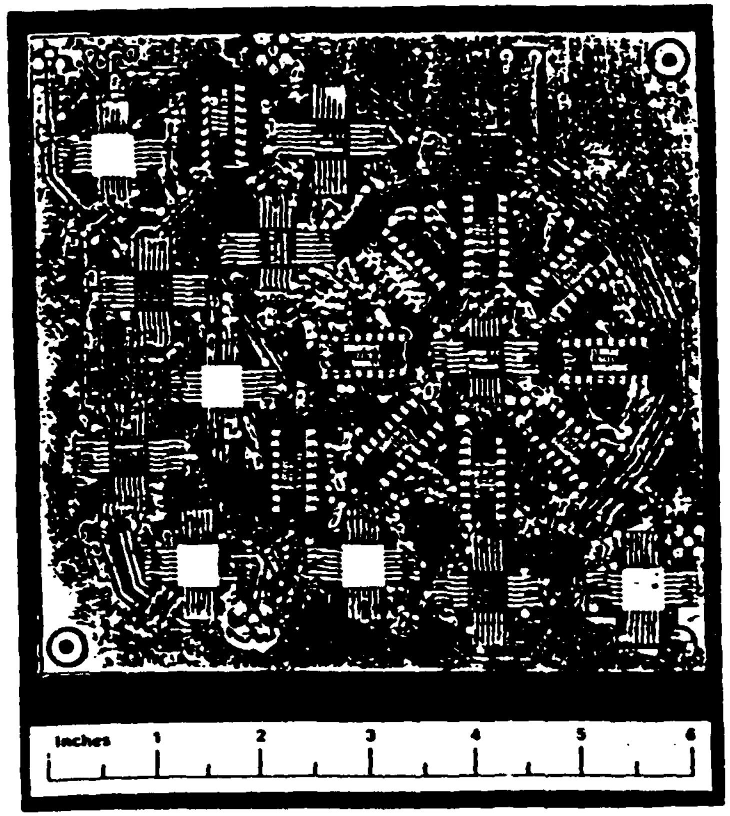


FIGURE 9 PHOTOGRAPH OF 256-BIT PN GENERATOR BOARD

SLIDE 28

SO MUCH FOR EXCUSES..........

QUESTIONS ANSWERED

SLIDE 29

HOW FAR CAN SPREAD SPECTRUM SYSTEMS COMMUNICATE THROUGH FOLIAGE?

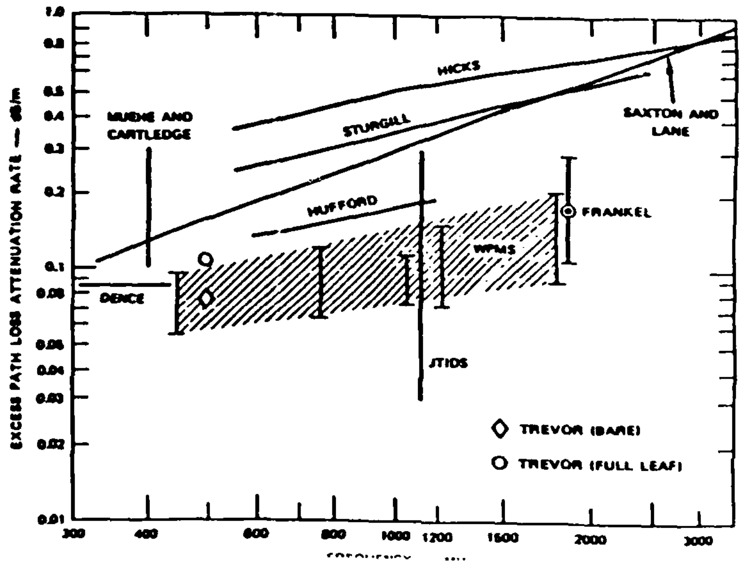
PATH LOSS IS FREQUENCY AND POLARIZATION DEPENDENT

PATH LOSS IS STRONGLY DEPENDENT ON TREE DENSITY (IN TRUNK REGION)

LEAVES CAUSED 5-6 DB INCREASE IN ATTENUATION THROUGH DECIDUOUS CANOPY

(SEE FIGURE 3.1)

DO NARROWBAND PREDICTIONS APPLY TO SPREAD SPECTRUM?



SLIDE 30

WHAT BANDWIDTHS ARE SUPPORTABLE IN FOLIAGE?

WHAT IS THE COHERENCE BANDWIDTH THROUGH THE FOREST?

MULTIPATH DELAY SPREADS OF HUNDREDS OF NANOSECONDS WERE MEASURED, UP TO AS MUCH AS 1 MICROSECOND OR EVEN MORE IN EXCEPTIONAL CIRCUMSTANCES.

IF COHERENCE BANDWIDTH IS $\frac{1}{\text{[redacted] (Delay Spread)}}$

THEN COHERENCE BANDWIDTHS OF <10 MHZ ARE TYPICAL

SLIDE 31

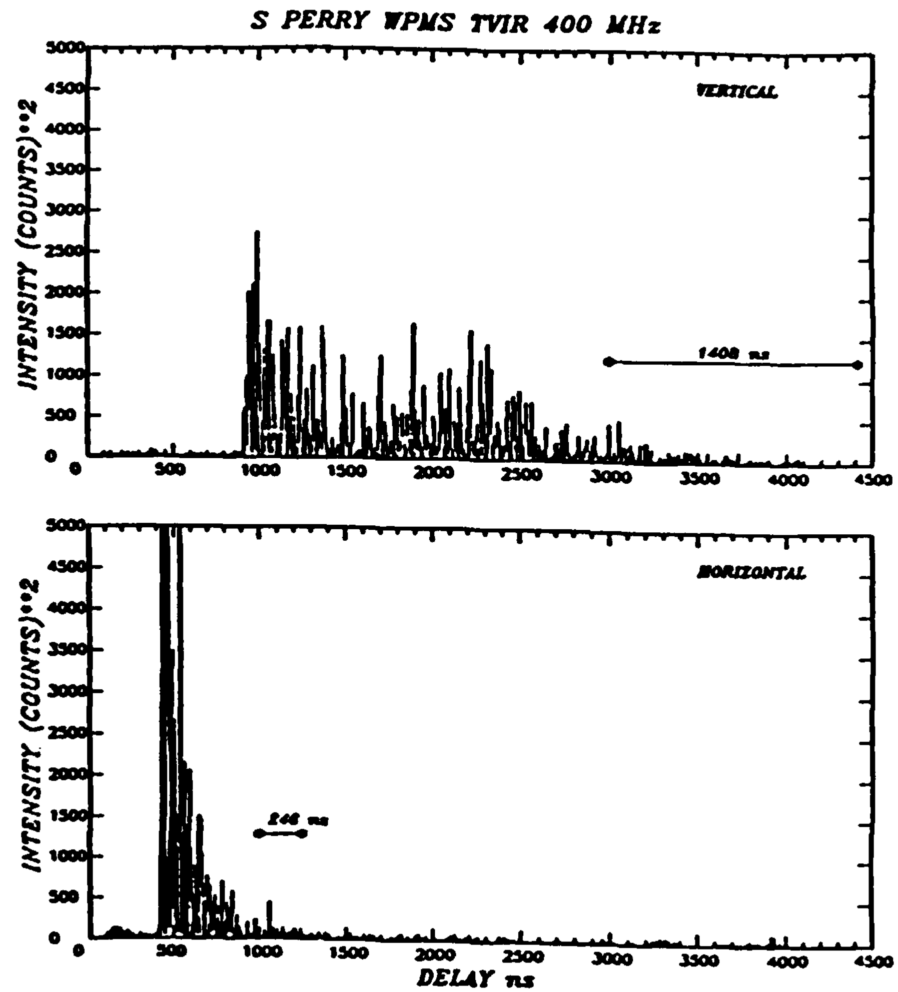


Figure 3.8 South Perry Woods Power TVIRs Measured at 400 MHz with Characterization Level and Calculated Delay Spread Indicated

SLIDE 32

frequencies (below 1 GHz) and at the higher chip rates.

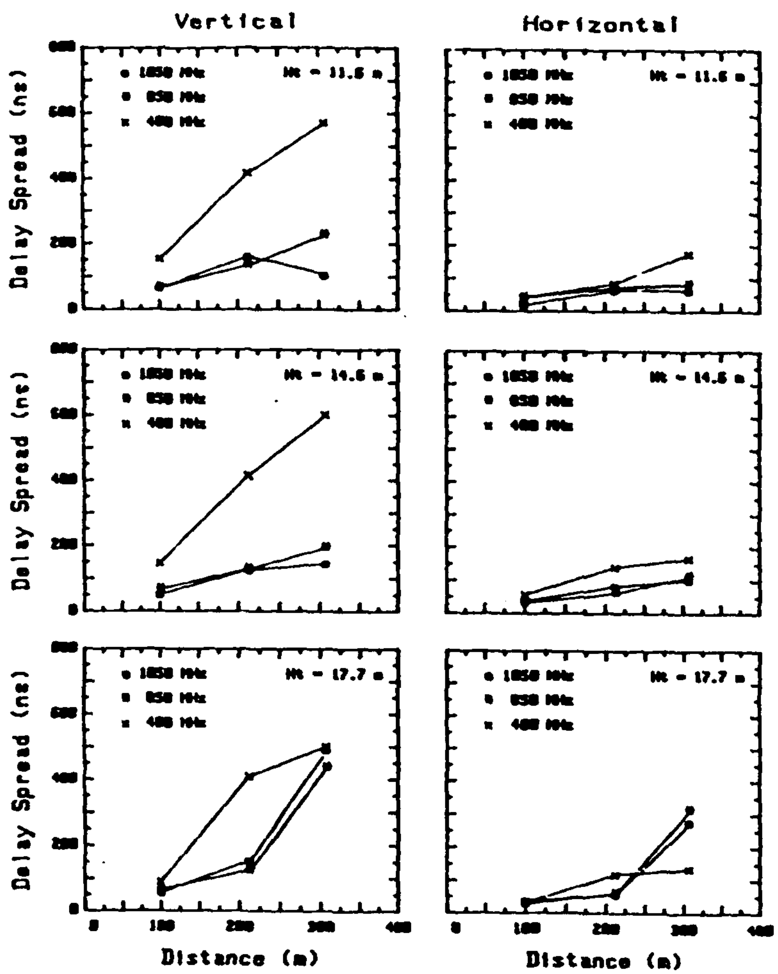
O.K., I'm going to very quickly review some of the answers we've gotten. Al Schneider is going to talk some more about the detailed work. The first question we were asking is: Just how far can spread spectrum systems communicate through the trees? We found a lot of data, much of which we haven't even yet analyzed, but we've found very definite frequency and polarization dependencies. Al Schneider has since developed models which have been reasonably well validated, particularly in the trunk region of the forest. So we are fairly comfortable with our ability to predict forest trunk region performance of UHF signals. We didn't spend much time in the leaves in Connecticut. We ran out of time and money but we did find that propagation through the leaf canopy also adds on the order of 5-6 dB of attenuation. Of course I think he'd be very unhappy with me making a blanket statement like that.

Another question relates to how well the narrowband predictions made in the past apply to spread spectrum. Here we found categorically, across the board, that spread spectrum predictions need to account for all the scattered energy. Unless you account for all that scattered energy received as in our wideband received signal level, you're obviously throwing out energy that exists at the spread spectrum receiver. The narrowband measurements, all of which are shown here, don't typically account for any of that scattered energy. Therefore their measured path-loss is significantly higher than what we measured in the shaded region using integration of all the received scattering energy. Therefore, as you would expect, our conclusion is that you cannot rely on narrowband measurements to predict received signal level of spread spectrum systems.

Another question we asked had to do with the existence of resolvable multipath in the foliage. The answer here is no. Although much of the measured TVIR data appears to indicate resolvable multipath, Al Schneider has completed more comprehensive analysis of the multipath returns and confirmed that increasing bandwidth continues to produce new multipath components. Delay spread measurements are reasonably representative, however, typically producing coherence bandwidths of less than 10 MHz. I showed you what the typical delay-spread looks like in the forest. One of the things that we definitely found was that delay-spreads with vertically polarized antennas were almost always significantly greater than delay spreads with horizontally polarized antennas. Al's model also predicts that in the trunk-dominated forest and intuition kind of confirms that.

We got interesting results in the delay spread measurements and this SLIDE 32 confirms what I said; basically that delay spreads with vertically polarized antennas (on the left) were significantly more in all cases than those with horizontal. We've measured a variety of distances, a variety of frequencies, all shown here, and a variety of antenna heights (as you go up the page) all below the canopy. Towards the top we were starting to see some possible canopy effects, as we approached the canopy. So we saw some non-ideal behavior in the data.

Another question related to antenna heights: [SLIDE 34] Where you put your antennas and what kind of antennas you use in the trees. In the trunk region we saw very little height gain (or what's been called height gain) attributed to the antennas. It really didn't matter what the antenna height was, as long as you were down in the trunks. Path-



DELAY SPREAD MEASUREMENTS IN TRUNK DOMINATED FORESTS FORT LEWIS, WA

SLIDE 33

WHAT ARE THE BEST ANTENNA HEIGHTS/ POLARIZATIONS/ ANTENNA PATTERNS FOR COMMUNICATIONS IN FORESTED ENVIRONMENTS?

LITTLE HEIGHT DEPENDENCE OF PATH LOSS OR DELAY SPREAD IN TRUNK REGION

PATH LOSS USUALLY GREATER FOR VERTICAL POLARIZATION THAN HORIZONTAL

DELAY SPREAD USUALLY GREATER FOR VERTICAL POLARIZATION

DELAY SPREAD DECREASES AS TREE DENSITY DECREASES IN TRUNK REGION

IN DECIDUOUS CANOPY ANTENNA HEIGHT HELPS WITH VERTICAL POLARIZATION BUT NOT WITH HORIZONTAL

DELAY SPREAD INCREASES WITH RANGE, ESPECIALLY FOR VERTICAL POLARIZATION

(SEE FIGURES 3.3, 3.6)

SLIDE 34

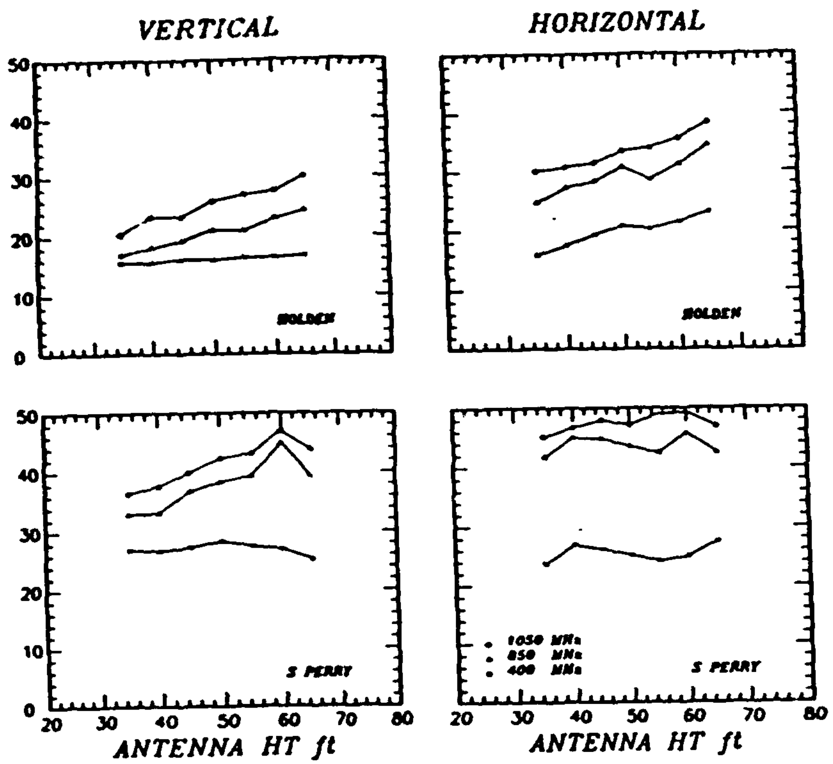


Figure 3.3 Comparison of 1986 Height-gain Patterns on Comparable Holden and South Perry Woods Paths

SLIDE 35

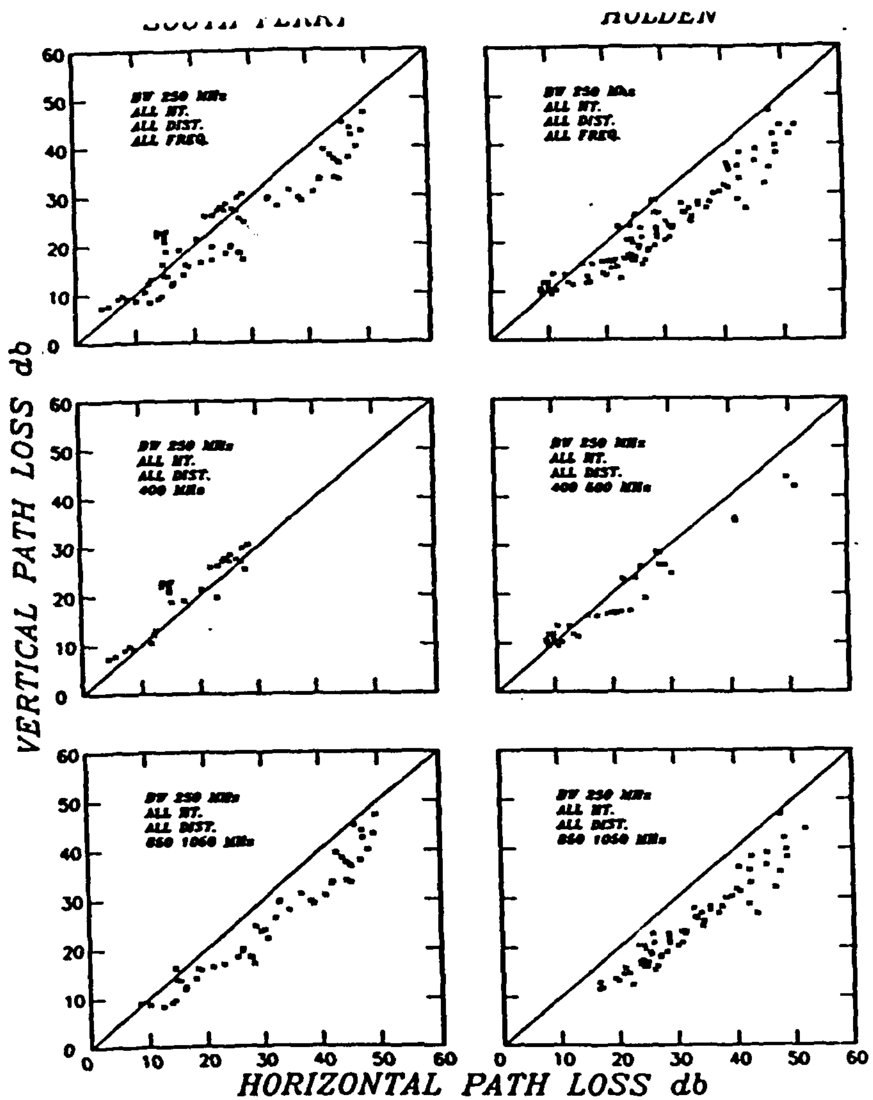


Figure 3.6 Summary of 1986 Polarization Dependencies for Tower Measurements

SLIDE 36

WHAT IS THE UHF PROPAGATION MECHANISM IN FORESTS?

INCOHERENT SCATTERING WITHIN TRUNK-DOMINATED REGION

NO EVIDENCE OF LATERAL WAVE

<u>SLIDE 37</u>

WHAT PHYSICAL FOREST PARAMETERS DETERMINE FOREST COMMUNICATIONS CAPABILITY?

TREE DENSITY (STEM COUNT PER UNIT AREA) IS THE CONTROLLING PARAMETER AFFECTING BOTH PATH LOSS AND MULTIPATH DELAY SPREAD.

(SEE FIGURES 3.13 AND 3.14)

<u>SLIDE 38</u>

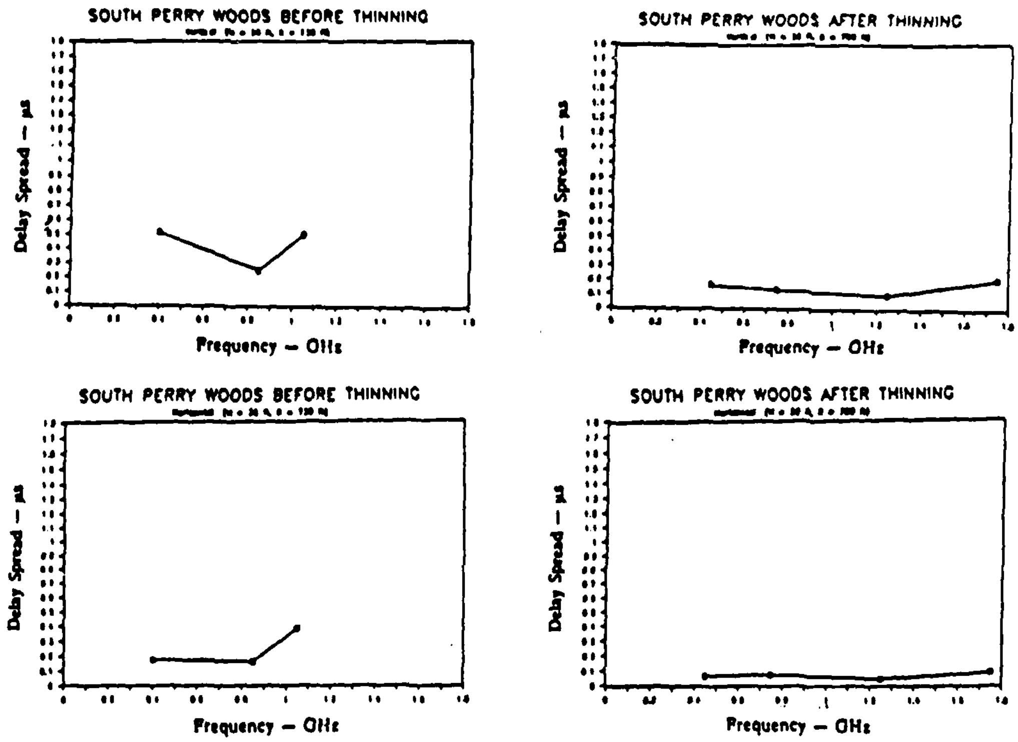
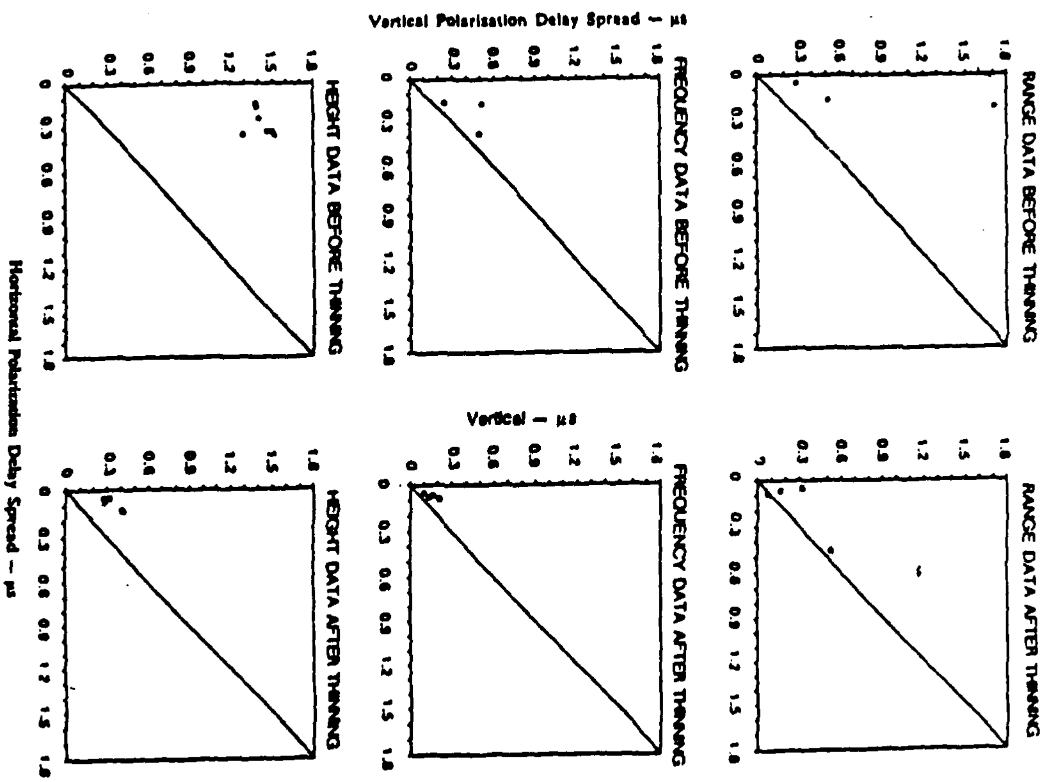


Figure 3.13 Multipath Delay Spread Versus Frequency Before and After Thinning

SLIDE 39



SLIDE 40

WHAT EFFECTS DO LEAF AND GROUND MOISTURE, GROUND CONDUCTIVITY HAVE?

NONE NOTICED , EVEN AFTER RAIN AT BOTH FT. LEWIS, WA AND COVENTRY, CN.

WHAT TIME VARIABILITY IS OBSERVED?

TIME INVARIANT, EXCEPT FOR TERMINAL MOTION

SLIDE 41

loss was usually greater for vertical polarization than for horizontal polarization, and delay-spread was greater for vertical polarizations. As expected, we saw a very definite dependence of delay-spread on tree density. We had a unique opportunity to conduct measurements in the same exact forest before and after it was thinned at Fort Lewis, and that provided some very unique and interesting data. I think the curve I want to rely on appears in a few slides. We went to a forest that had been planned for cutting by the foresters at Fort Lewis and we persuaded them to let us conduct measurements both before and after they made a 25% reduction in the density of the trees. The results were very interesting.

As far as the propagation mechanism, there had been some conjecture by a number of people that there exists a so-called "lateral wave" which was an indirect up, over and down mechanism, almost a wave-guiding effect along the top surface of the trees. We're fairly comfortable now concluding that we haven't seen any evidence of a "lateral wave" at these frequency ranges. The lateral wave is a mechanism that was fairly well accepted in the 60s and 70s at VHF frequency ranges and it was being extrapolated with very little grounds up to higher frequencies. We don't feel that is correct. My one frustration is that we didn't quite obtain the path lengths we would have liked to in the forest. The forest was just too lossy. And, as I said, the incoherent scattering within the trunks was very dominant.

As far as the forest parameters that determine communications capability, just about the most dominant one, by far, was tree density. The number of trees per square area (per acre) or whatever your measurement is, is the most important and sensitive parameter. Tree type, seasonal variations, and moisture had no observable effect. Next to tree density, the size of the trees (i.e. the width of the trunk) had a measurable effect.

This is the curve I was alluding to before. It shows a return to the Fort Lewis forest before and after the thinning. A lot of the data shows this general behavior, but we unfortunately don't have a lot of data points. But the trends are generally fairly noticeable and we have to go back to the raw data to extract more in-depth measurements. Returning to the same trees after the thinning significantly reduced the delay spread both for horizontal and vertical polarizations. [Question on leaves.] Again, this applies in the trunk region. And these were coniferous trees, so there were no leaves, only needles.

As far as stationarity of the channel, we saw no time-variance at all in any of our measurements unless the terminal was moving. There had been some reports of leaves blowing, changing the measured spectrum. In fact, I've used them before in a paper I wrote almost 10 years ago ... leaves changing the visible spectrum in a JTIDS signal. We didn't see anything like that. We didn't see any variability with time, although there admittedly wasn't very much wind. We didn't look at very many leaf-channels in high wind conditions, so that might be the reason. We saw no effect of leaf and ground moisture, ground conductivity, or any of those parameters that the people that study trees tend to claim will effect propagation. We didn't see it.

My last slide is simply a plea to this community [SLIDE 43]. We know all the theoretical applications of measurement data. We heard about some this morning. Nevertheless, I'm really at a loss to propose how we now use this data. I'd like to see indications of interest from this community. I would like

APPLICATIONS FOR CHANNEL CHARACTERIZATION DATA

- THEORETICAL RECEIVER PERFORMANCE EVAL
  - CANONICAL FSK/ PSK MODEMS
  - RAKE/ PDI/ CRI
  - ADAPTIVE DECISION FEEDBACK EQUAL
  - MAX LIKELIHOOD SEQUENTIAL ESTIMATION
  - DIVERSITY COMBINING
- CHANNEL SIMULATORS
  - SYNTHESIZED CHANNEL
  - STORED CHANNEL
- EMPIRICAL/ THEORETICAL PROPAGATION MODELS
  - LONGLEY RICE PROPAGATION MODEL
  - FOREST SCATTER MODEL
- PROPAGATION MECHANISM INTERPRETATION
  - LATERAL WAVE EXISTENCE
- COMMUNICATIONS SYSTEM DESIGN
  - DELAY SPREAD $\Longrightarrow$ IRREDUCIBLE ERROR RATE

SLIDE 42

THE $64,000 QUESTION

WHAT DO WE STILL NEED TO KNOW ABOUT COMMUNICATIONS CHANNELS?

HOW CAN WE GET THE NECESSARY DATA?

SLIDE 43

to offer any of the data, any of the reports in any form you'd like, to any of you to use for any of your work. This was, as I said, a long exercise and we've made a lot of progress for the forest environment. I don't think we've made much progress in a non-forested environment and that, in a large sense, is the channel that I think, motivated Mike Pursley's comments. I know that recently he's been talking about the delay spread giving a good measure of the irreducible error rate in the channel. I'd like to see how that can be exploited with the data we've acquired.

LINDSEY: Thank you Paul. We have one final presentor, Al Schneider, regard to measurements of data.

ALLAN SCHNEIDER: *Delay Spread Estimation for Time Invariant Random Media*

The topic of this talk, its title anyway, is "Delay-Spread Estimation for Time-Invariant Media." I've broadened the topic somewhat because of some of the things Paul Sass indicated earlier in his talk. I thought you might be interested in seeing some of the results to date of our modeling and characterization of this forest channel. First, I'm going to give you an overview of the incoherent channel forest model, and then after that I'll get into more of the modelling details, and, finally, identify one of the problems that we had in analyzing the data. One of the problems really gave us a lot of pain until we finally discovered one solution for it. I'm not saying our solution is optimal, but at least it gave us one way of getting consistent results. Of course, one of the reasons I'm here is to solicit help. You people here who have more experience in data processing and hypothesis testing than we, might be able to see alternative ways so that we can improve the analysis of this data.

In channel modeling or radio wave propagation, there are things that I have become very sensitive to over the years. First of all, people that sponsor radio wave propagation studies always ask themselves when it's over, if they have thrown their money down some rat hole. Propagation studies do not come cheaply as you saw by this one - three-and-one-half million dollars, roughly. And the question is, have we gotten anything for our money? So, right up front, when we started working on this project, one of the things that I did was to take the set of questions that Paul had laid on the team [which included SRI, Penn. State and CyberCom], and to keep that posted in front of my desk to remind me that we were always working toward some practical objectives and practical problems. The second thing, to temper that study, was to ask ourselves, what could we reasonably expect out of some sort of channel model or propagation model. For those of you who have some background experience in tropospheric scatter, for example, the Booker-Gordon theory - I don't know when it was proposed, some time around 1955 or so - has existed for a very long time and there has been a lot of analysis done on it and so forth. But when all's said and done and you're faced with the practical problem of designing a tropospheric scatter circuit, the last place you really want to start is with the Booker-Gordon scattering formulation. Because it simply does not provide quantitative answers that are consistent with measured data. It does a very good job in giving you some clue as to what the sensitivity of a certain parameters is, but, quantitatively, when estimating, say, the transmission loss, whether it is 170 dB or 110 dB, Booker-Gordon is not where you want to go. With that in mind, I felt that we had to keep our feet on the floor if we were to recognize what

can we reasonably expect from a model that is significantly much more involved than the scattering in the troposphere. Scattering in the troposphere, for example, involves electromagnetic scattering in random media but it's single scattering. Scattering in the forest involves multiple scatter, a much more difficult problem.

So with that prologue as preface, let me get into the forest model as a modeller might. One of the first things you might want to concern yourself with is, what are the environmental forest parameters that are likely to affect the transmission loss, delay- and Doppler-spread of a forest channel [SLIDE 1]. We tried to identify the key parameters that might affect forest propagation by representing the forest as a planar, stratified medium consisting essentially of the forest floor, then above the forest floor a trunk-dominated region, and above the trunk-dominated region a canopy consisting of branches and leaves. It was our immediate objective to model this random medium by using discrete scatterer theory - the Foldy-Lax multiple scattering theory. You can find a description of it in Ishimaru's two-volume text (Academic Press, 1978). It allows us to account for polarization effects as well as multiple scattering. Our model is, essentially, the Foldy-Lax theory. The way it works is, you assume that the medium consists of a very large number of discrete scatterers, and each of these scatterers is characterized in terms of its so-called T-matrix, or, equivalently, in terms of its dyadic scattering amplitude. The dyadic scattering amplitudes are, in some sense, similar to the dyadic Green's function that Bill mentioned earlier. The forest components (tree trunks, branches, and leaves) are represented by so-called canonical scatterers. Specifically, trunks were modelled as dielectric cylinders having some prescribed complex dielectric constant; branches the same way; and leaves were modelled as disks. These canonical scatterers can be characterized by their so-called dyadic scattering amplitudes - you take a single scatterer, put it out in free space, and describe the scattered far-field over the full 4π solid angle in response to an incident plane wave. In the case of tree trunks, you then assume that all the tree trunks are standing parallel to each other and perpendicular to the forest floor. They are randomly positioned on the forest floor, they have prescribed distribution of trunk diameters, but they are all parallel to each other. In the canopy we assume that the branches are arbitrarily oriented with prescribed, azimuthal and elevation angle statistics; the leaves have prescribed angle statistics. We then take the dyadic scattering amplitudes and tumble-average them over the angular distributions of the scatterers in the canopy. On the basis of these models, we were able to incorporate most of these forest parameters. For example, with the tree trunks, certainly the key parameter, as Paul Sass mentioned previously, is the tree trunk number density, (the number of scatterers per hectare of forest floor per acre), but also important is the tree trunk diameter probability density function and also the effective height of the trunks to where the canopy begins. And you can imagine how many other parameters might bear on this problem - the forest homogeneity (tree trunk number density varying from place to place), the reflection coefficient of the forest floor, the tree water content, the dielectric constant of green wood. All these parameters, and perhaps others, might enter into our characterization of the forest channel.

Based on the Foldy-Lax theory, we were able to characterize the scattered field. It

UNIVERSITY OF SOUTHERN CALIFORNIA

COMMUNICATION SCIENCES INSTITUTE

WORKSHOP ON

ADVANCED COMMUNICATION PROCESSING TECHNIQUES

DELAY-SPREAD ESTIMATION FOR TIME-INVARIANT RANDOM MEDIA

PRESENTED BY

ALLAN SCHNEIDER

MAY 1989

CYBERCOM CORPORATION

4105 NORTH FAIRFAX DRIVE

ARLINGTON, VIRGINIA 22203

[TEL: (703)-522-0046]

CyberCom

FOREST MEASUREMENTS	FOREST PARAMETERS
o TRUNK NUMBER DENSITY	o FOREST COMPOSITION
o TRUNK DIAMETER PROB. DENS. F'N	o FOREST AGE
o CANOPY HEIGHT & THICKNESS	o FOREST BASAL AREA
o BRANCH NUMBER DENSITY	o LEAF AREA INDEX
o BRANCH DIAMETER PROB. DENS. F'N	o SEASON (FULL LEAF, NO LEAF)
o BRANCH LENGTH PROB. DENS. F'N	o FOREST HOMOGENEITY
o BRANCH EULER ANGLE PROB. DENS. F'N	o FOREST FLOOR
o LEAF NUMBER DENSITY	o FOREST WETNESS
o LEAF AREA	o TREE WATER CONTENT
o LEAF THICKNESS	o PERMITTIVITY OF WOOD & LEAVES
o LEAF EULER ANGLE PROB. DENS. F'N	

SLIDE 1

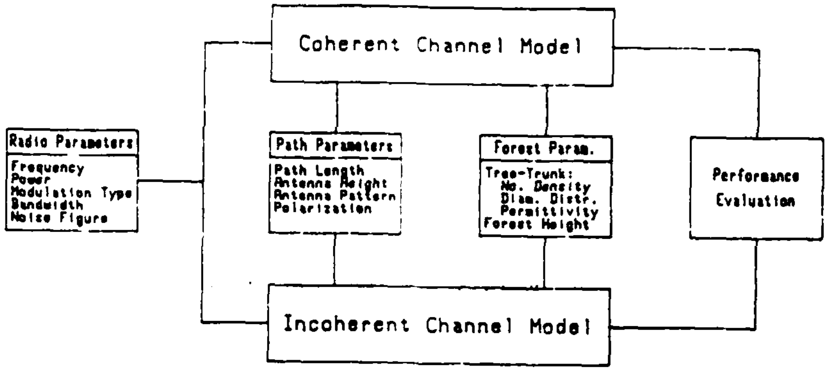
turns out that if the scatterers are small, in terms of a wave length, they scatter isotropically. And when they scatter isotropically, the Foldy-Lax theory allows you to get an exact solution for the mean scattered field and also an equation for its correlation function. For the space-frequency correlation function, you can get a correlation equation for which you can get an exact solution when the tree trunk diameters are small in comparison to the wavelength. When you go to larger tree trunk diameters, the Foldy-Lax theory becomes very involved – computationally tedious – and then you have to go to something like transport theory, in which case you don't get exact solutions. You get big numerical solutions. We used a hybrid technique. We used the Foldy-Lax for the smaller tree trunk diameters, and then we used transport theory for the larger ones. Using an approximate technique, we also extended the thin-trunk Foldy-Lax model to the thick trunk case. I can't go into exactly how we did that here, but I'll be glad to discuss it afterward.

On this basis, we developed the following model for the forest [SLIDE 2]. We also assumed that the key radio parameters were the frequency, of course, the transmitter power, the modulation type, the bandwidth, and the noise figure. Ultimately we were hoping to get some measure of signal-to-noise ratio and also some measure of performance in a frequency dispersive channel. The forest channel itself was divided into two parts. One was called the coherent channel model and the other one the incoherent channel model. This is a common dichotomy, but an unfortunate choice in terminology especially for communications people because communicators like to measure coherence on the basis of phase stability. For electromagneticists, their idea of the coherent component is really the ensemble average; the fluctuation about the ensemble average is what constitutes the incoherent channel component. So there's a question of semantics here which is fundamental to understanding the characterization of this channel. And it's especially so in this case because the forest, for all intents and purposes, is time-invariant. At least over this frequency band where the frequencies range from, let's say, 200 MHz to 2000 MHZ, and the wavelengths range from roughly 1 meter to 10 cms. In order to effect any substantial forest movement, you have to move one of the scatterers an appreciable portion of the wavelength. Since the tree trunks tend to remain fairly well fixed (planted if you'll excuse the pun), there's very little time fluctuation and phase incoherence. In other words, once you transmit a signal, because the channel is time-invariant, the received phase of that signal does not vary with time.

When you have a large number of randomly positioned scatterers, it turns out that the average field drops rapidly to zero, except at very low frequencies. In the VHF band, for example, the coherent channel model is the dominant one and gives rise to the lateral wave investigated by Tamir. But at the higher frequencies, it's the incoherent channel model which is the most important one, and that's the only one that I'm going to discuss today.

The model equations themselves are horrendously complex as you might expect, dealing with complex quantities, dyadic scatterers, random media, and all the other things that go up to make the forest channel. Nevertheless, it is possible to get engineering models [SLIDE 3] that people can actually use in practical situations and that's what I'd like to show you in the next few viewgraphs.

SLIDE 4 presents the output menu of the



SLIDE 2

Figure 1-2: Forest Channel Model

CyberCom Incoherent Forest Transmission Model

for

Specific Attenuation and Delay Spread

SLIDE 3

computer program. It comes up on the screen when you push RUN. What I'm going to show you is how you run the model, enter the input parameters, and obtain the outputs. Afterward I'll get into the nitty-gritty of how we actually try to measure some of these delay spread functions. The first thing that the model asks you is "What do you want in terms of an output?" We provide three different options: one is specific attenuation, the second one is albedo and cross-sections, and the third one is delay spread. From the communications point of view it's the total transmission loss which is of primary interest, but in most practical cases the forest is nonhomogeneous and since the specific attenuation is directly proportional to the number of trees per acre, you really have to integrate the specific attenuation over the path profile of the trunk number density in order to come up with the estimate of the transmission loss on a non-homogeneous forest. So the most fundamental quantity in terms of transmission loss really turns out to be the specific attenuation and that's usually measured in terms of dB per meter. The other output option that's of major interest to communicators is delay spread. We want to know how the signal spreads in time - what its dispersive characteristics are - because we want to know such things as intersymbol interference and what we can expect from intersymbol interference in terms of performance (bit-error rates). The second output option (albedo and cross-section) is more of a intermediate output. It helps the user to interpret some of the results of the options 1 and 3. The albedo, just to refresh your memories, is the ratio of the scattering cross-section of a target to the total cross section. This output option plots that ratio, the albedo, and also the cross sections of the scatterers themselves. These quantities depend on the size of the scatterers and also on the frequency.

Certainly the most important forest parameter is the tree trunk number density (the number of trees per acre), but second is the tree trunk diameter probability distribution [SLIDE 5]. Now in practical terms, you can't expect soldiers to go out through the forest and measure tree trunk diameter distributions each time they want to set up a tactical radio system. That doesn't make much practical sense. And if we had to come up with the tree trunk diameter distributions for forests throughout the world, again the model would be virtually unusable. It turns out that the foresters have found out that there are essentially two generic classes of tree trunk diameter distributions. One is the so-called even-aged forest. An even-aged forest is one where perhaps through natural calamity, all the trees have been burned off - like Yellowstone, two years ago - or like in a clear cutting case where the foresters go in and harvest the whole area of trees. The forest is completely wiped out, seedlings sprout up and after ten years or so you have an even-aged forest. But all trees don't grow to the same size - not all grow as rapidly - and there is a distribution of tree trunk diameters; this diameter-distribution tends to be normal. Alternatively in a natural forest - one that has matured and is continuing to renew itself - the diameter distribution turns out to be exponential. Foresters often call it the "inverse-J" distribution, but it is the same as the exponential distribution. These diameter distributions constitute the so-called canonical forests that the model is based on. You specify whether it is even-aged or uneven-aged in order to use one of those models. Alternatively you can enter the specific tree trunk diameter distributions you want.

```
Which one of these do you want to plot?

1- Specific Attenuation   2- Albedo/Cross-Section
3- Delay Spread           4- None of the above
-----------------------------------------------


Type 1 to 4, then (ENTER)
```

SLIDE 4

```
Enter Forest Type:

1-Even-aged   2-Uneven-aged   3-Measured   4-South Perry   5-Coventry
---------------------------------------------------------------------
Trees in the 0 - 2  inch bin =  0
Trees in the 2 - 4  inch bin =  30
Trees in the 4 - 6  inch bin =  76
Trees in the 6 - 8  inch bin =  82
Trees in the 8 - 10  inch bin =  77
Trees in the 10 - 12  inch bin =  67
Trees in the 12 - 14  inch bin =  52
Trees in the 14 - 16  inch bin =  37
Trees in the 16 - 18  inch bin =  25
Trees in the 18 - 20  inch bin =  11
Trees in the 20 - 22  inch bin =  11
Trees in the 22 - 24  inch bin =  1

Total number of trees          =  469

Are these data correct (Y or N)?

Enter trunk diameter truncation limits


Truncation limits are  0 - 24  inches

Do you want to plot the Histogram of Trunk Diameters?--(Y or N)

Enter Trunk Density (stems/ha) [Default=1000]

Enter Path Length (m) [Default=305]
```

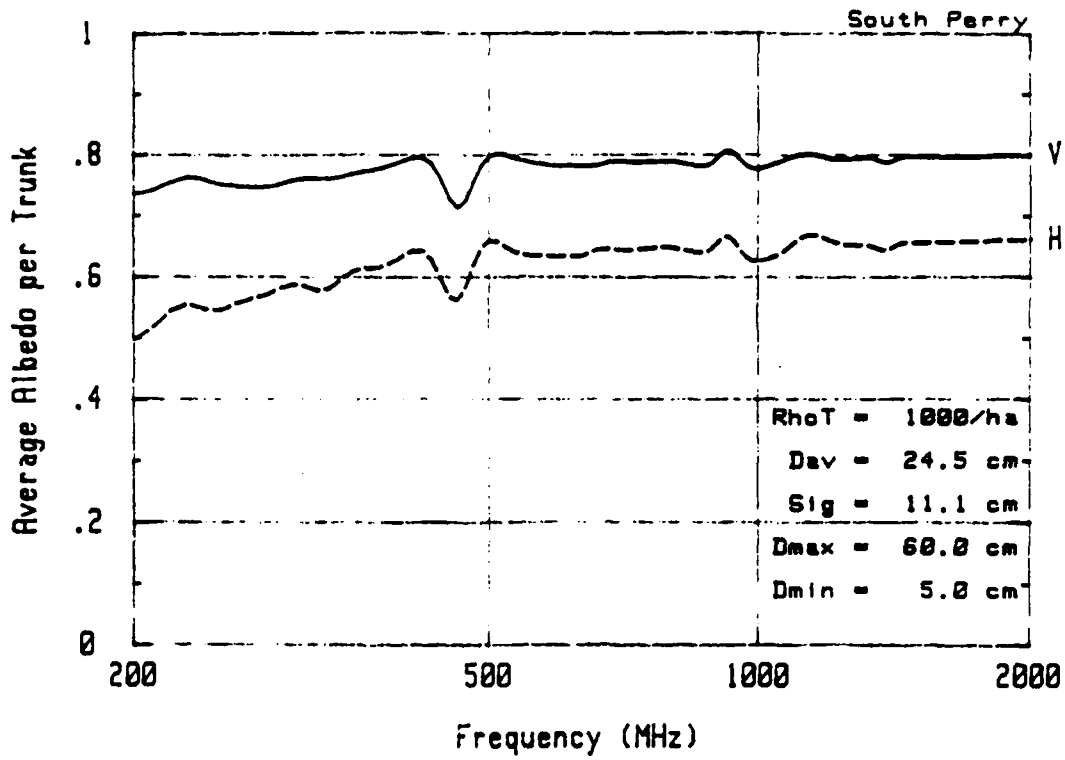
SLIDE 5

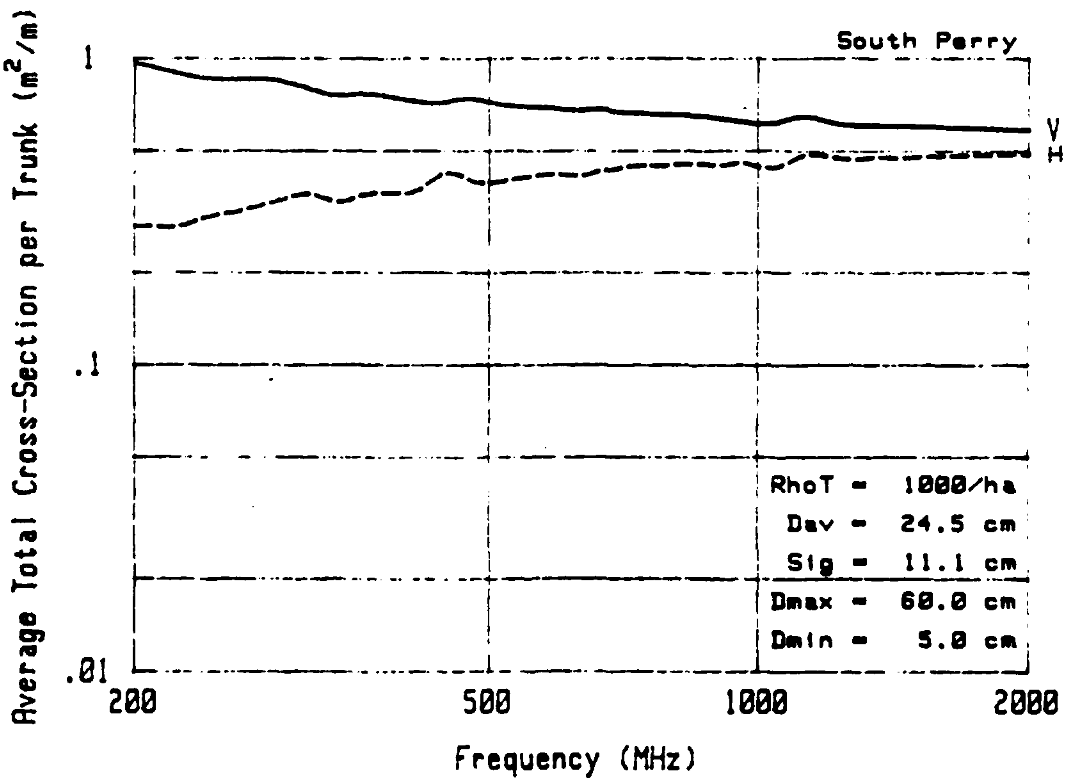
With that information, knowing the parameters of the forest, you are able to start getting some of the output [SLIDE 6] (I want to pass over the intermediate calculations quickly). I call your attention, in the first case, to the upper curve, that plots the average albedo per tree trunk, based on the tree-trunk diameter probability distribution. You get different albedos depending on polarization – you can see roughly that it is about 0.6. That's a key parameter to keep in mind when I show you the next slide. The total cross section is shown lower down in the viewgraph.

Perhaps the most important parameter in any radio communication system is transmission loss. If you don't have enough signal you're not going to get any reliable communications whether you have delay spread or not. This plot [SLIDE 7] shows the variation of the specific attenuation in dB per meter as a function of frequency and corresponds to the South Perry tract where we made the measurements. It's normalized to 1000 trunks per hectare (a hectare is a 100 meters on each side), the average diameter is 24.5 cms (they were rather husky trees), and the standard deviation and other parameters of the distribution, are identified in the figure. The upper curve represents the attenuation of the so-called coherent component while the bottom curve represents the attenuation of the incoherent component. It's the incoherent component which dominates – it has less specific attenuation than the coherent does. You can see at the highest frequencies for example, the coherent specific attenuation is somewhat greater than 0.20 dB per meter where the incoherent components it is probably 0.15 By the time you integrate over a 100 meter path, the coherent component becomes negligible.

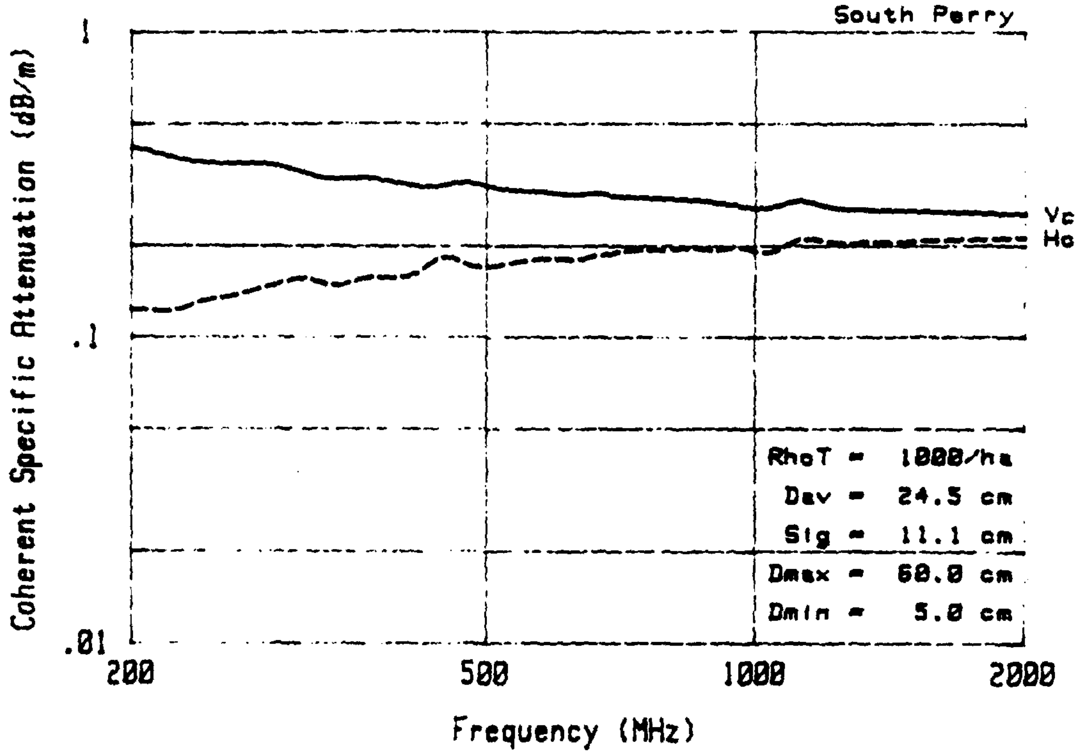
One thing that you may notice from this is first of all, the ratio of the two attenuations is approximately equal to the albedo – about .6 – or it's inverse, three-halves or 1.5. And the second thing we can notice from these curves is that as the frequency increases, the polarization dependence of the channel model essentially disappears. This is consistent with theory which predicts that as objects become much larger in terms of a wavelength, they tend to lose their polarization dependence. So what you would do is, if you want to find out the transmission loss on a particular path through the forest, you'd run the model, come up with the specific attenuation appropriate to that particular tree trunk number density and diameter distribution,and then multiply it by the path length in that portion of the forest. Obviously, it's a numerical form of WKB integration.

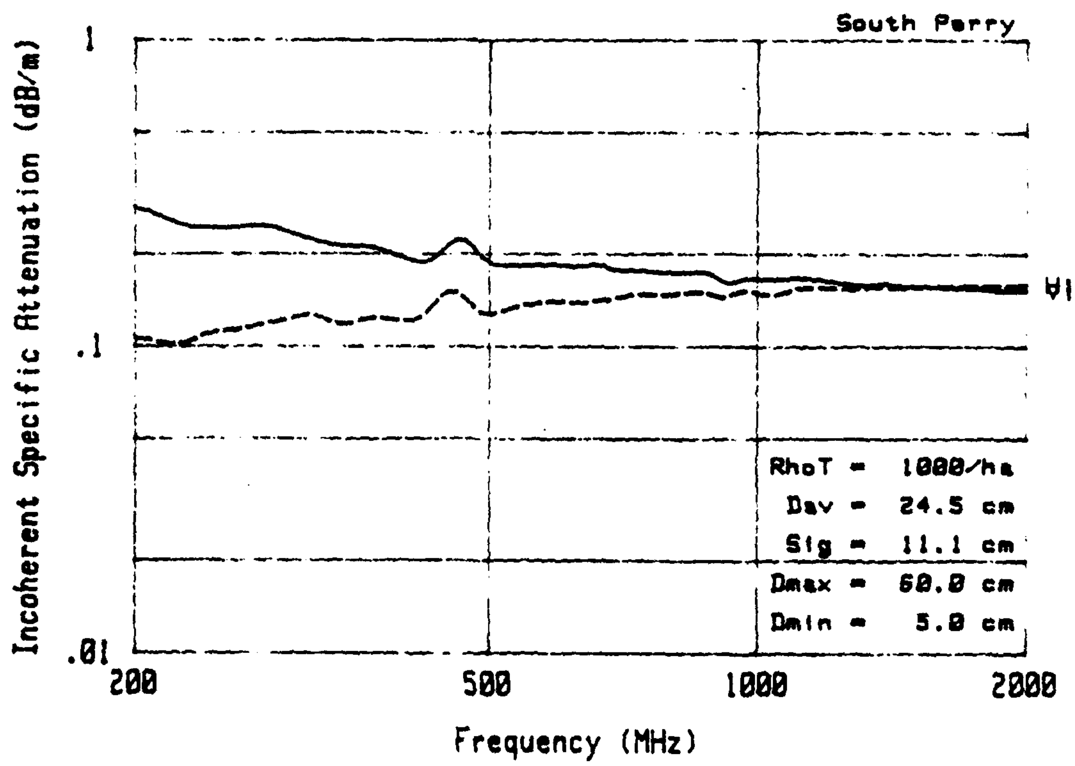
The second thing that the model allows you to do is to compute the delay spread. The delay spread is nominally a function of distance – how far you propagate through the forest – but it turns out that in an unbounded random medium – which is what we have here (we assume that the tree trunks extend an infinite distance laterally) – the delay spread reaches some asymptote and levels off with no increase. That's what this curve shows [SLIDE 8]. These are the asymptotic limits. It shows that the delay spread for vertically polarized waves is significantly higher than it is for horizontally polarized waves. You'll notice that there are occasional minima, for example, near 450 MHz. They can be correlated with creeping waves propagating around the circumference of the trees. If, for example, you find out what the circumference of the tree is and compare it with the wavelength, you'll find out that they correspond to resonances of creeping waves around the tree trunks. In practice, how important they are,


South Perry
Average Albedo per Trunk
1
.8
.6
.4
.2
0
V
H
RhoT = 1000/ha
Dav = 24.5 cm
Sig = 11.1 cm
Dmax = 60.0 cm
Dmin = 5.0 cm
200
500
1000
2000
Frequency (MHz)

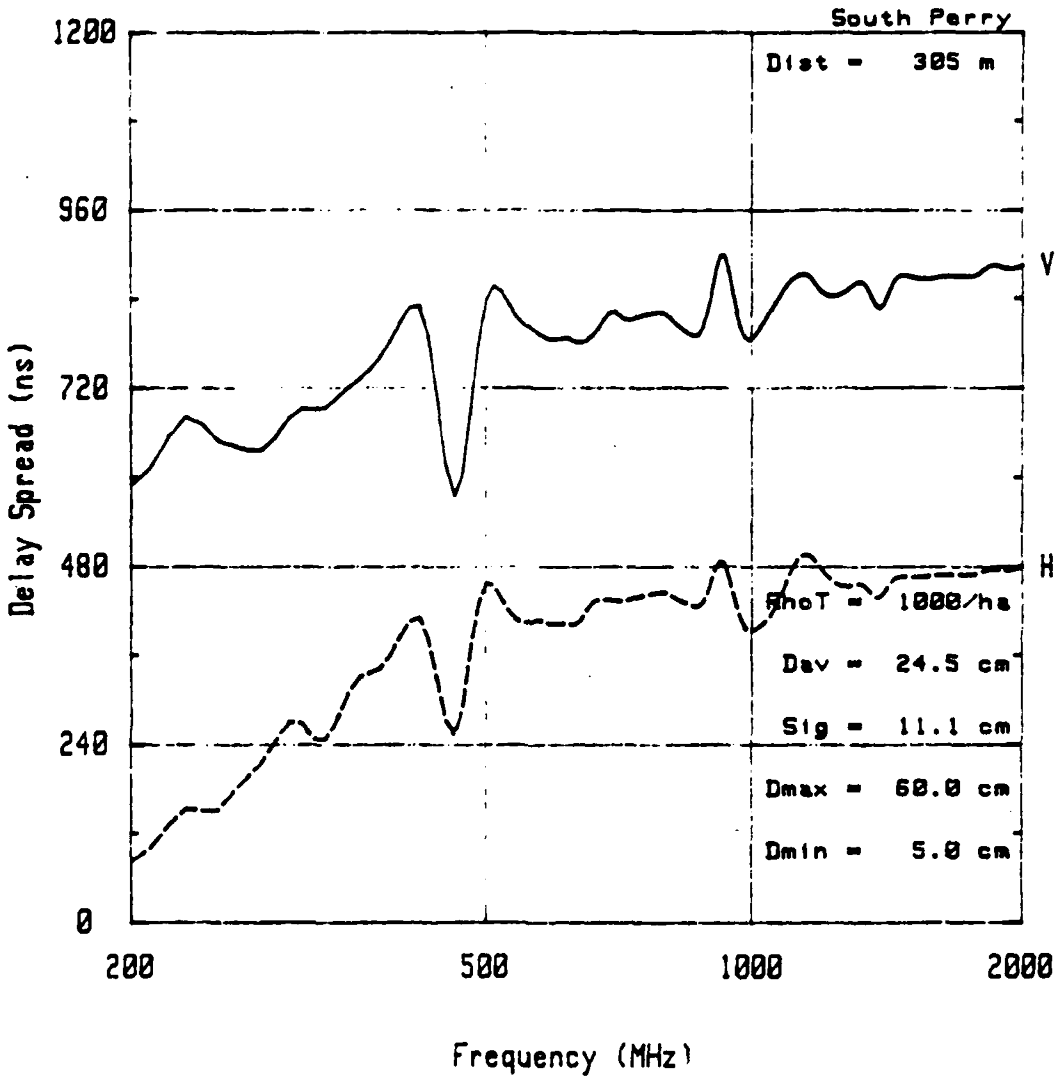

South Perry
Average Total Cross-Section per Trunk (m²/m)
1
.1
.01
V
H
RhoT = 1000/ha
Dav = 24.5 cm
Sig = 11.1 cm
Dmax = 60.0 cm
Dmin = 5.0 cm
200
500
1000
2000
Frequency (MHz)

SLIDE 6


South Perry
1
.1
.01
200
500
1000
2000
Frequency (MHz)
Coherent Specific Attenuation (dB/m)
Vc
Hc
RhoT = 1000/ha
Dav = 24.5 cm
Sig = 11.1 cm
Dmax = 60.0 cm
Dmin = 5.0 cm


South Perry
1
.1
.01
200
500
1000
2000
Frequency (MHz)
Incoherent Specific Attenuation (dB/m)
RhoT = 1000/ha
Dav = 24.5 cm
Sig = 11.1 cm
Dmax = 60.0 cm
Dmin = 5.0 cm

SLIDE 7


South Perry
Dist = 305 m
RhoT = 1000/hs
Dav = 24.5 cm
Sig = 11.1 cm
Dmax = 60.0 cm
Dmin = 5.0 cm
1200
960
720
480
240
0
Delay Spread (ns)
V
H
200
500
1000
2000
Frequency (MHz)

SLIDE 8

I don't know because the tree trunks are not perfectly smooth – there's bark and roughness there – *whereas we modeled the trees* as perfectly smooth, finitely conducting dielectric cylinders.

These then would be the parameters that an engineer might estimate from this particular forest channel model: total specific attenuation of the transmission loss, and delay spread to get the dispersion of the forest. Doppler does not arise because as I said, the trees, the leaves, etc. do not move appreciably in terms of a wavelength, and as long as the terminals are fixed, it means essentially that the forest is time-invariant – random but time-invariant.

If you want to see what a theoretical delay spread function looks like, it's plotted here [SLIDE 9]. I don't remember exactly what the forest parameters were. This is just a generic plot of it. It's essentially exponential but not quite. It actually has a little bit more energy around the origin than an exponential function might have. But for all intents and purposes, it looks like this.

Now the purpose of the measurement program, in addition to what Paul Sass said, was, from our perspective, not only to acquire measured data of delay spread and path attenuation, but to validate the model itself. You simply cannot afford to run 3 million dollar programs every time you want to find out what the specific attenuation or delay spread is in a forest of interest. If you want to extrapolate your measured data, you need a reasonable model; if you want to be confident that your model is a good model, you have to compare it with the measured data. And so that's what I'd like to address next and, as I do, I'm going to get into just one small problem that arose in the analysis of that data.

Before you go out to make a delay-spread measurement in the forest, this may be what the model predicts is the delay spread function [SLIDE 9]. When you actually go out in the woods to make the measurements, what you get looks something like this, a very ragged version [SLIDE 10]. Essentially what this raggedness reflects is the absence of ensemble averaging. In the model we've developed here, we have done ensemble averaging. We have taken forests of known parameters – for example, a tree trunk number density of 1000 trees per hectare with a prescribed diameter distribution – and have considered these trees randomly positioned on the forest floor; then we have taken essentially the same scatterers but reconfigured them on the forest floor, and repeated the theoretical modeling. We did it again and again, and took an ensemble average. By so doing, we got the smooth delay spread function you saw in the previous figure. But when you go out in the forest to make a measurement, what you see is what you get. You've got the trunks that are situated there. You go out and make your measurement and you get the delay spread function that is appropriate to that configuration – a sample from the total ensemble of all forests having the same parameters. It is not a simple case of saying, "OK, well, what I'll do to smooth this is, I'll simply move my antennas a little bit, or I'll go to another forest." When you go out and look at a forest, you may swear that forest is homogeneous spatially, but I'm telling you that it's not. God has made them all inhomogeneous. So that's what we're faced with – how can we do the two things that we'd like to do – estimate specific attenuation and delay spread – as we look at these two figures – the measured one, and the theoretical one? First of all, from the practical point of view of communications performance, we would like to

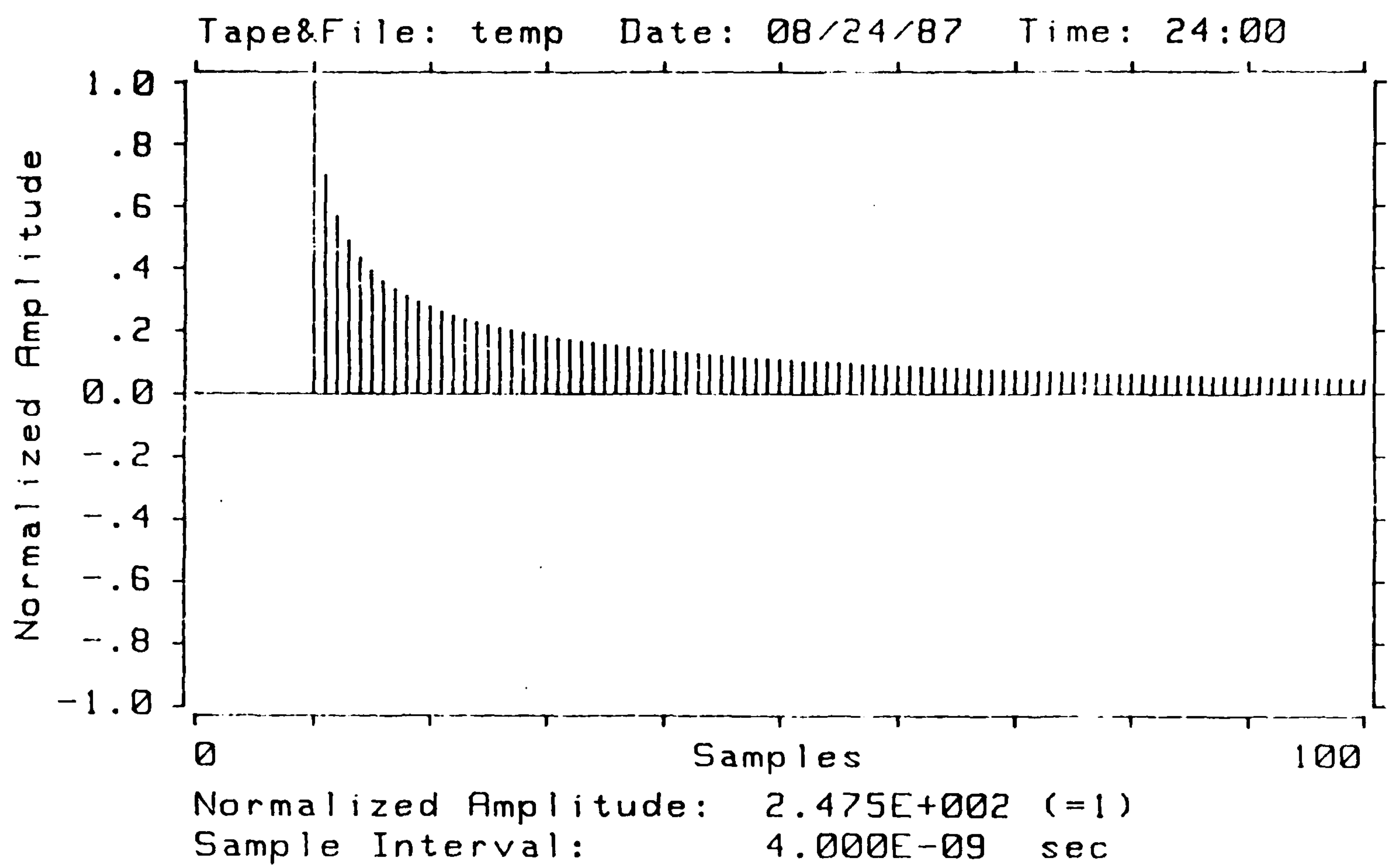
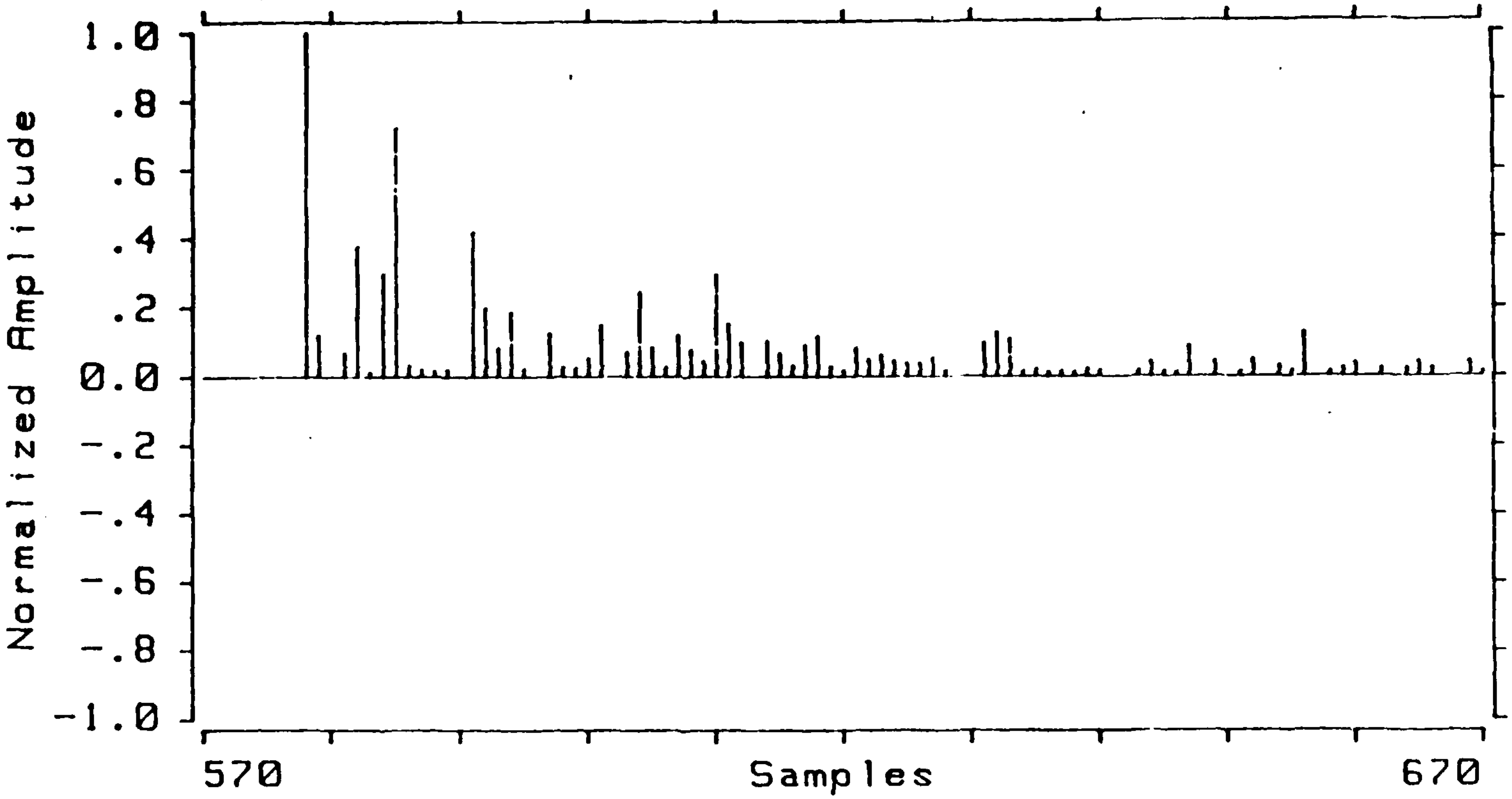


Figure 4-1: Delay-Spread Function (Ensemble Average)

SLIDE 9

Delay-Spread Function (Average)

Tape&File: T64F5R2 Date: 86/04/24 Time: 15:29



Normalized Amplitude: 1.053E+004 (=1)
Sample Interval: 4.000E-09 sec

SLIDE 10

be able to say that the delay spread has some measure of width, whether it be 3 dB width, or 1/e or 10 dB down – whatever we elect to use. Is the measure of delay spread the same for both the theoretical model and the measurement, or at least within reasonable agreement? And the second thing we would like to do – more or less from the point of view of confirming the validity of the model – is to determine if, in fact, this sample function is some sort of exponential. The model predicts essentially exponential dependence. Is this a sample function of a randomized sample of an exponential? Thus, there are two areas of interest here.

Well, the first thing we did, to give us a little bit of confidence, was ask, "OK, if this is what a model predicts on an ensemble average basis and if we can assume that at any given differential delay (we're plotting here received power as a function of differential delay), for example let's say this one right here, that the composite or ensemble average signal is actually arising from a large number of scatterers, it seems reasonable to a first approximation to assume that we essentially have the sum of a large number of vectors of random phase, and as we're plotting here the power, that would tend to be an exponentially distributed random variable." The voltage would be Rayleigh distributed, but the power would be exponentially distributed. We then took this theoretical ensemble average model [SLIDE 9] and we superimposed on it the Rayleigh fading, or if you will, the power exponential fading, and we got a function that looks like this [SLIDE 11]. If you compare the measured result which is the bottom one, and the model result which is the top one, they look reasonably good at least qualitatively. It's not a quantitative measure of goodness-of-fit but qualitatively they look pretty similar. But in order to get a quantitative measure, we had to get into the nitty gritty of actually assigning a measure of delay spread. The first approach that was taken was simply to use, as a measure of delay spread, the second central moment. It's a reasonable measure, it's convenient, its unambiguous, and you can compute it fast. The second central moment is also known as the standard deviation and should provide a good first estimate of the spread. But when we made those measurements and put these data into the computer to make the calculations, the values that we got for the second central moment turned out to be nearly an order of magnitude larger than one would suspect just from looking at the figure. We would look at the figure and say, "Well, delay spread looks like about 100 nanoseconds", and then would end up computing 700 nanoseconds. So when the data was plotted and correlated against different things like path length and frequency, the data points spread all over and looked inconsistent. So the questions were: what was causing that inconsistency, and what could we do about it?

That's what I'd like to take as my next topic – as long as Bill's not going to cut me off. Is that all right? ... [LAUGHTER]. Then let me be brief, and those that are interested in this problem, can see me afterwards.

Let me first of all describe what I had thought originally as the way I was going to fit this data. First of all I thought, "Let me test and see if it's an exponential." But just how would I test an exponential on this? The next thought was, "Well, I can simply fit some exponential curve that went through here using a least squares measure." But it turns out you really cannot do that because a least-squares measure does not provide an

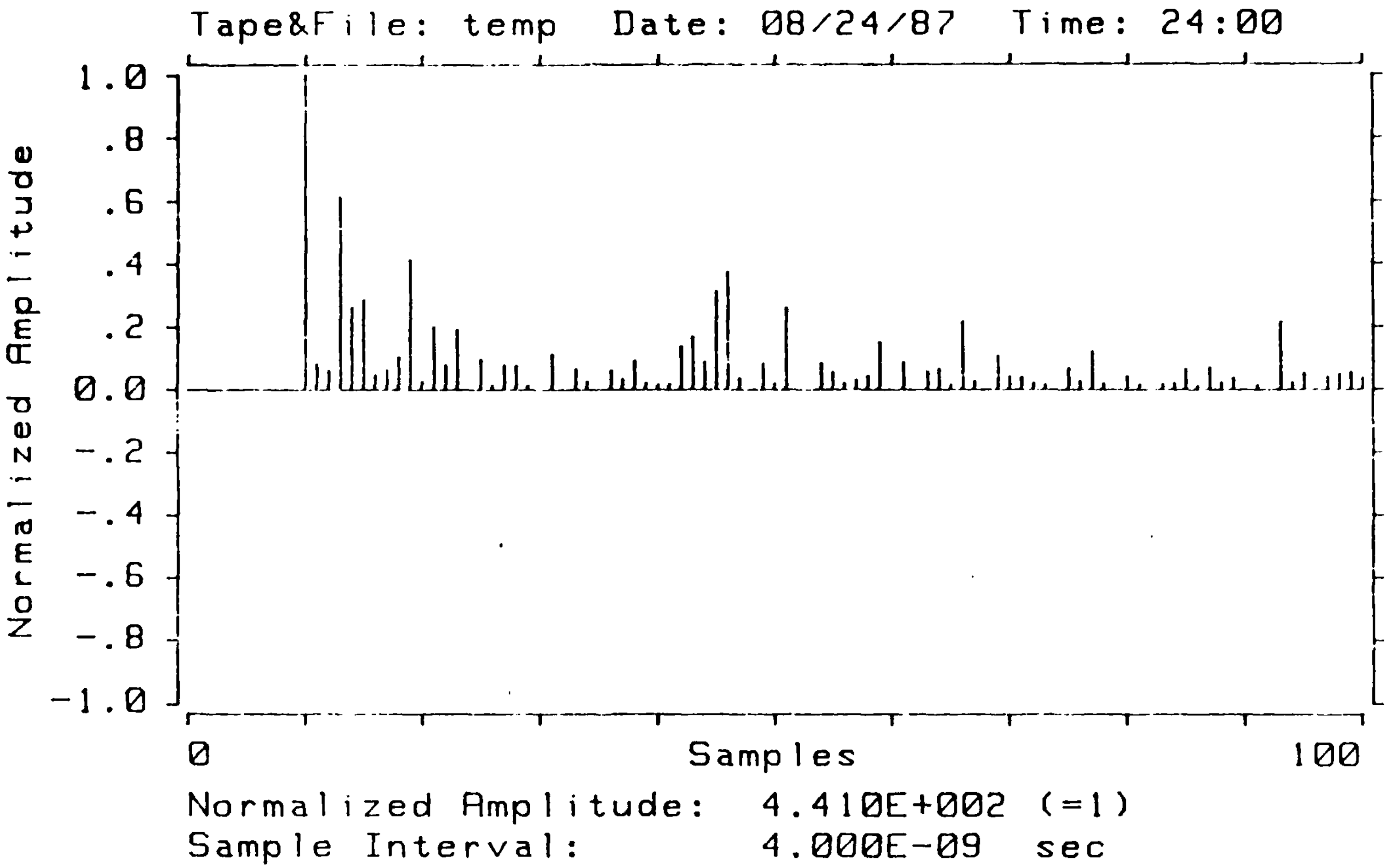


Figure 4-2: Delay-Spread Function (Sample)

SLIDE 11

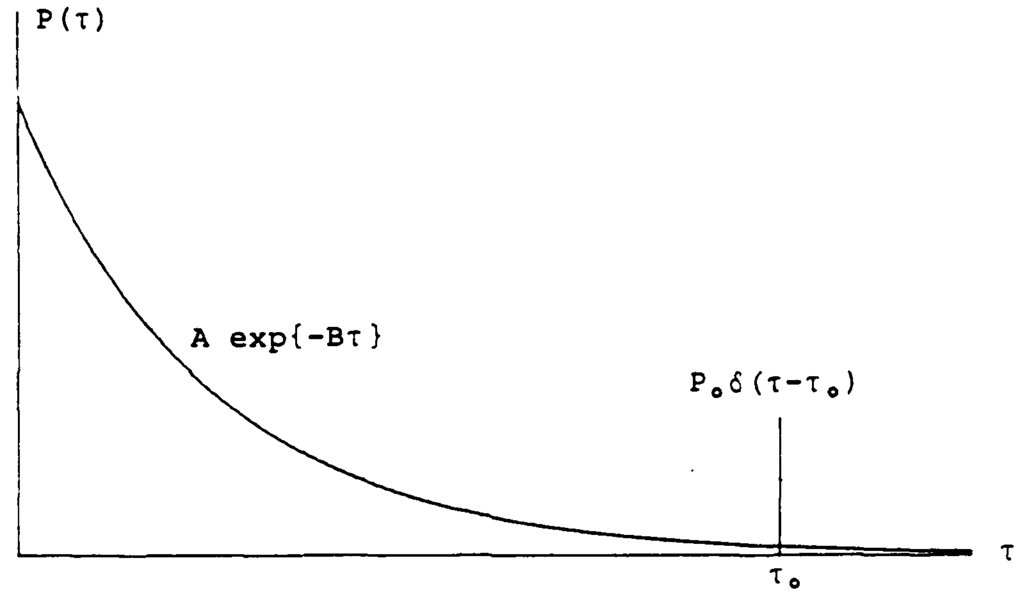


Figure 4-39: PTVIR with Outlier

SLIDE 12

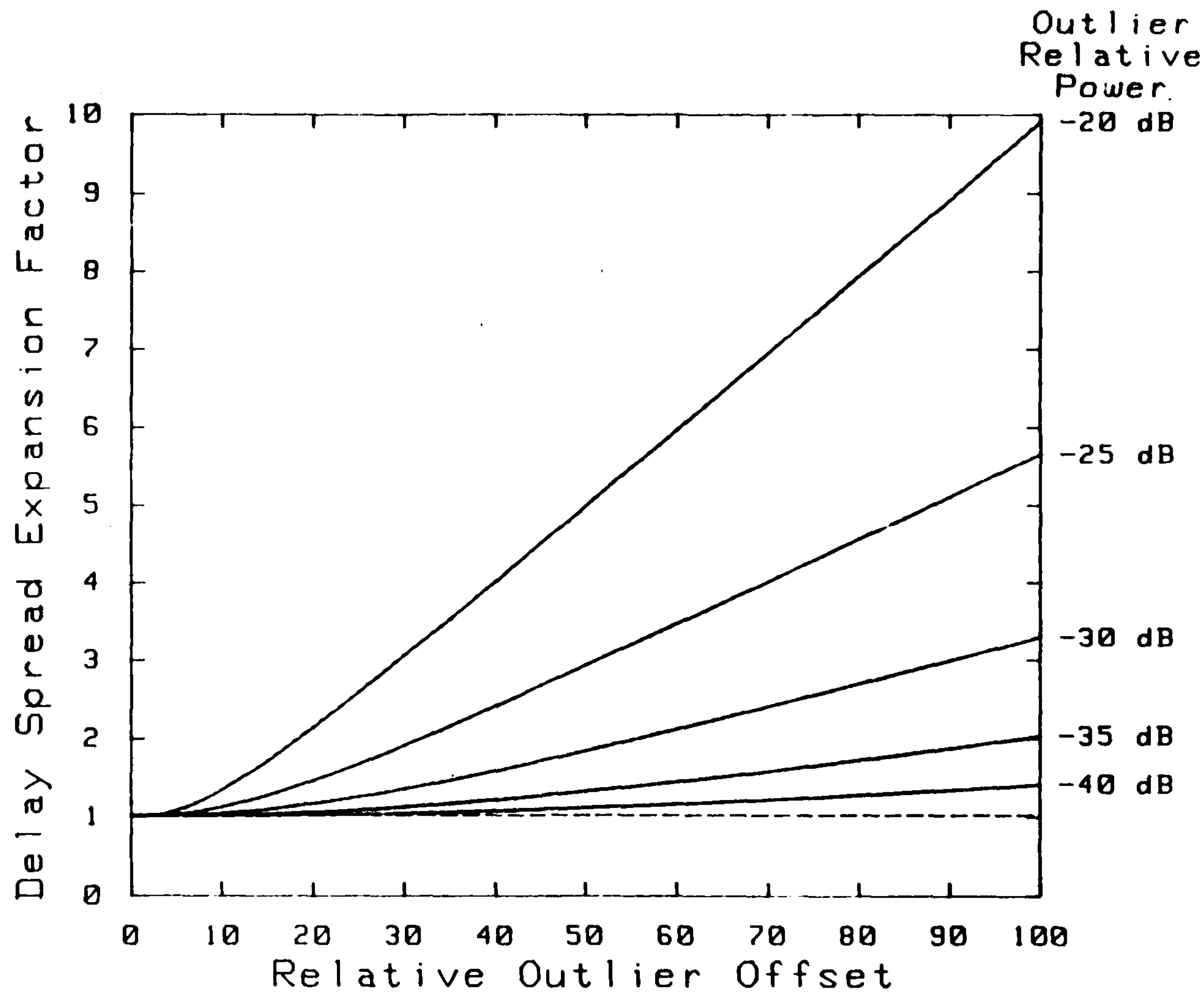


Figure 4-40: Delay Spread Outlier Expansion Factor

SLIDE 13

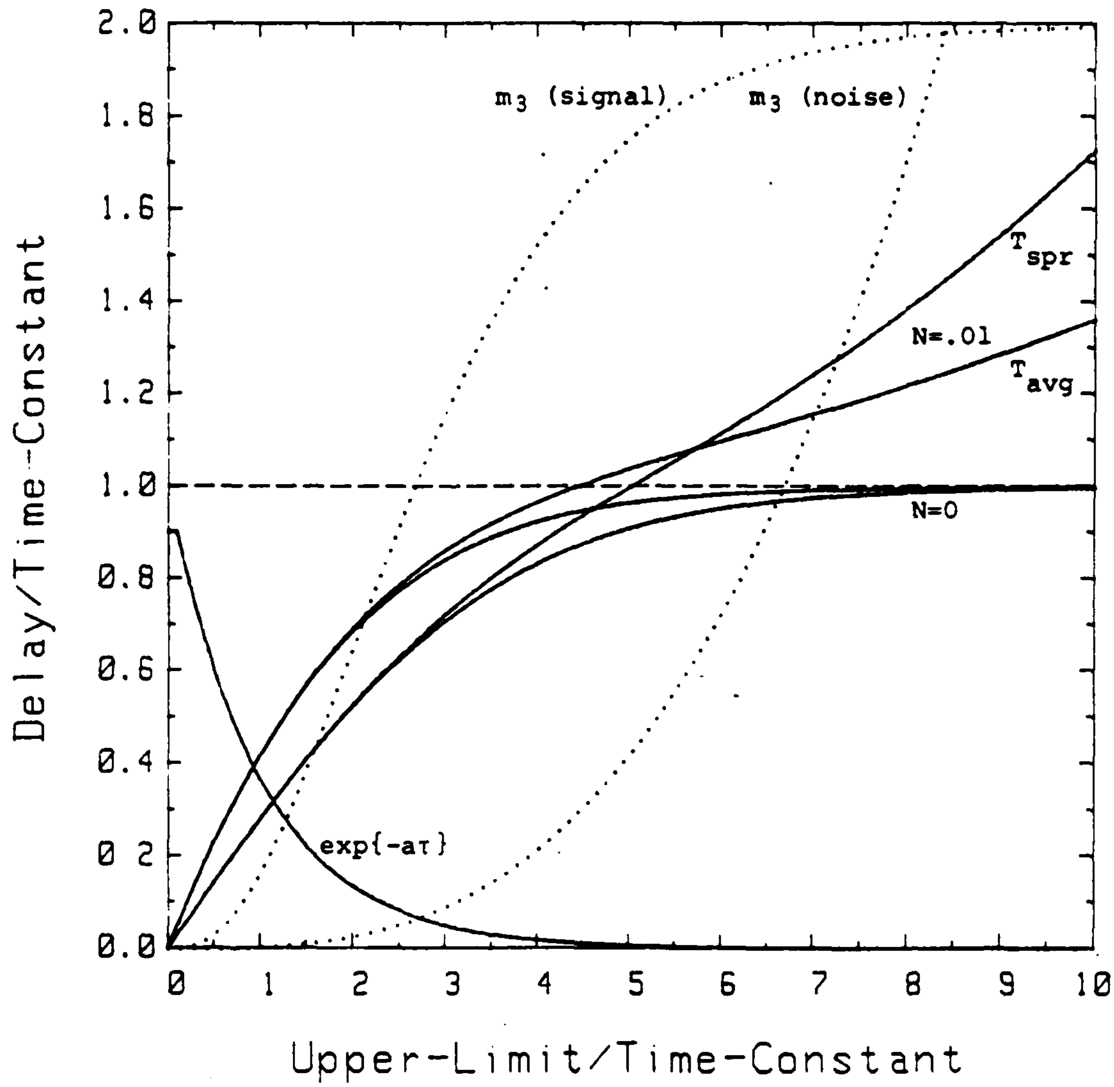


Figure 4-41: PTVIR Summation-Limit Sensitivity Curves

SLIDE 14

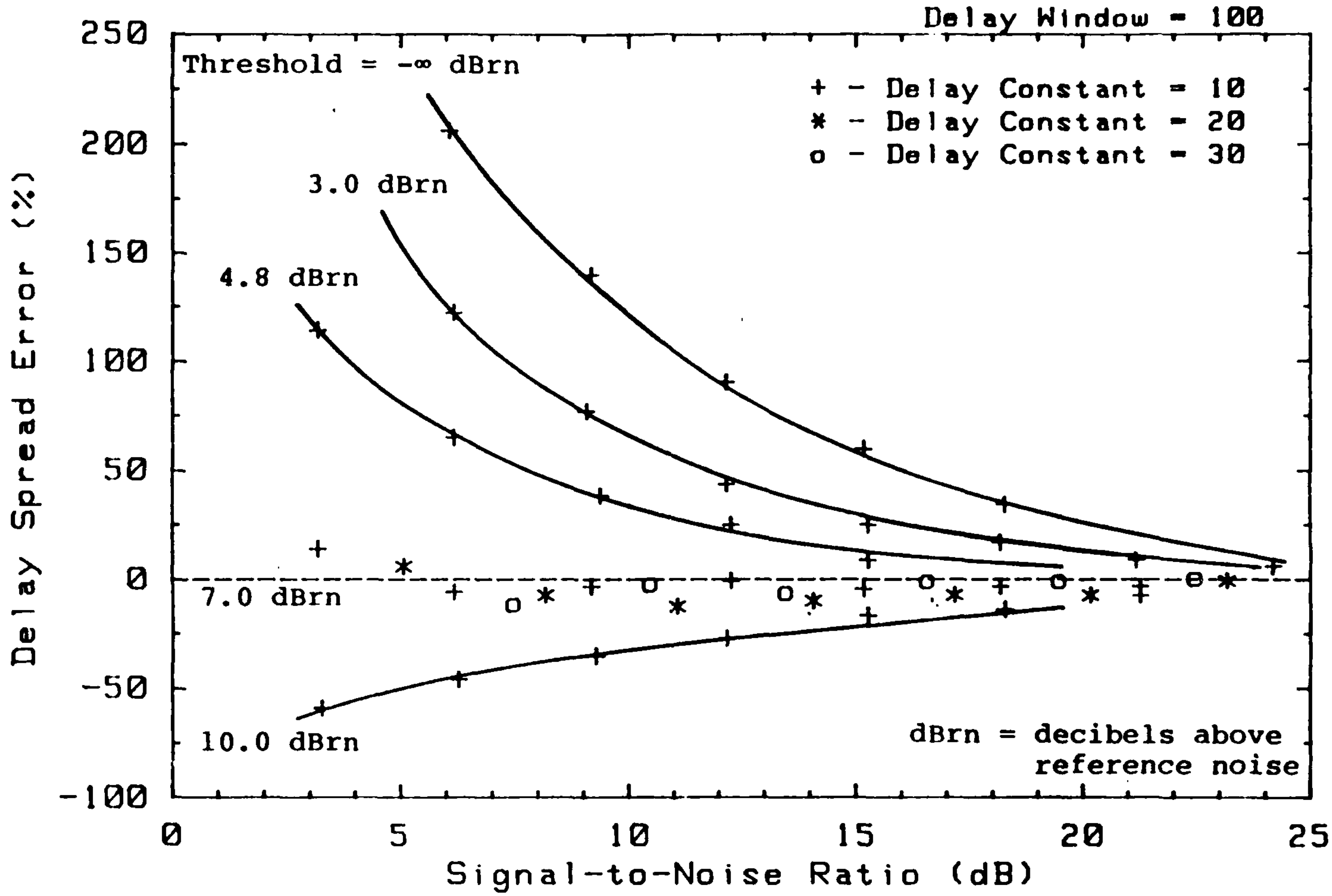


Figure 4-38: PTVIR Threshold Sensitivity Curves

SLIDE 15

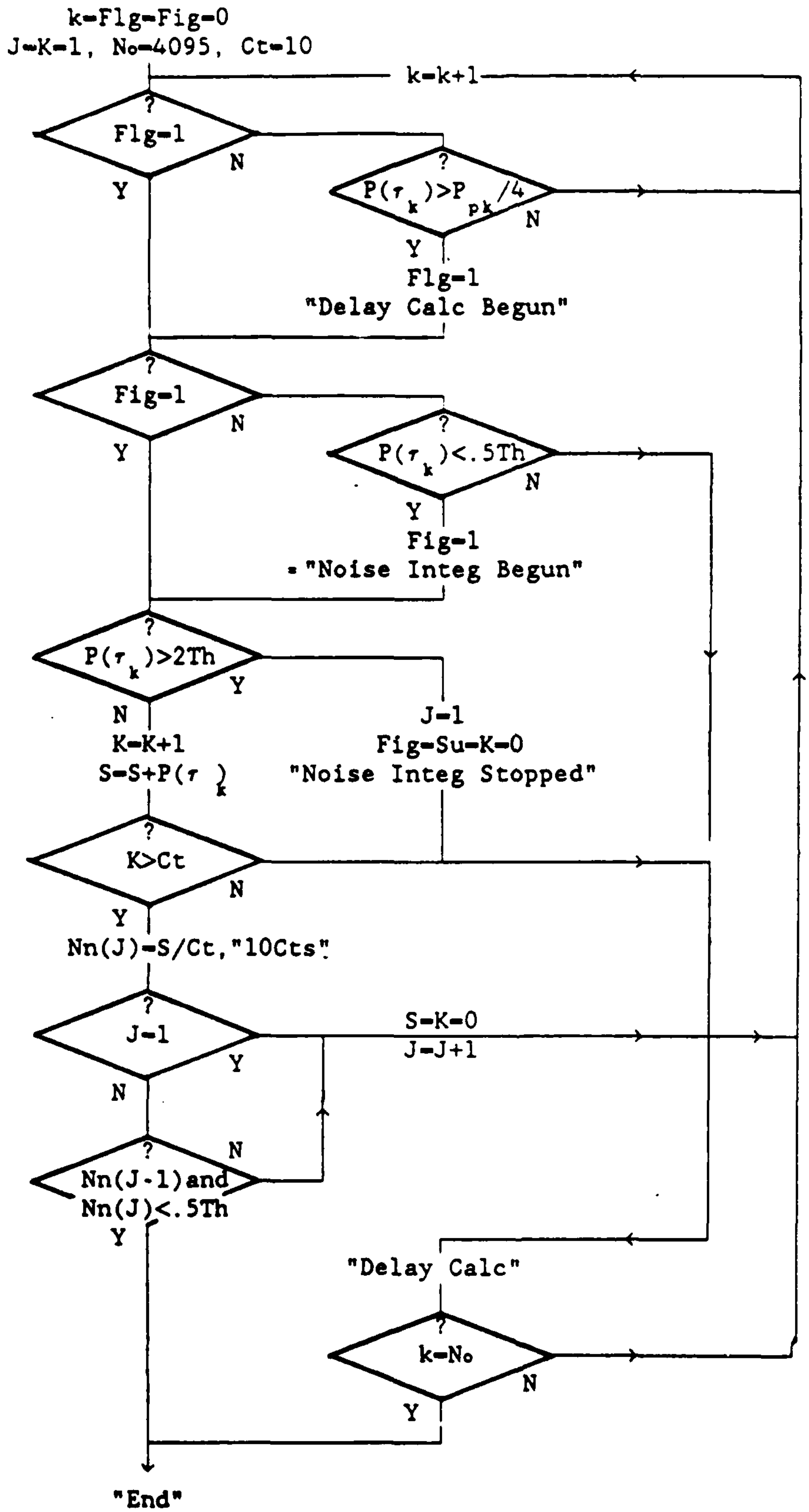


Figure 4-42: PTVIR Delay-Spread Algorithm

SLIDE 16

unbiased estimate – it turns out that Rayleigh fading tends to fade down more often than it fades up. It gives a bias and when I tried to employ that technique it did not work very well. Furthermore, I had to determine where the data began.

There were a lot of problems in trying to use the least-squares technique. So we began to look at what was causing the difficulties. Why were the calculated delay spreads so much larger than what appeared to be the correct answers? One of the things we looked at was what is the effect of an outlier on a delay spread calculation. This curve [SLIDE 12] represents the smooth exponential curve that might describe the general shape of a measured delay spread function. That little line out there represents some outlier that arises as a consequence of Rayleigh (exponential) upward power fading at a great delay. The second central moment depends on the square of the moment arm. As a consequence, even this isolated outlier can, if sufficiently distant with sufficient power, lead to all sorts of problems in overestimating the delay spread [SLIDE 13]. There were other problems as well [SLIDES 14 and 15]. For example, when you make a measurement of the delay spread function you're not measuring only the exponential-like curve representative of the scattered signal. In addition, you have background noise – far out here where there's virtually no signal – contributed not by the medium but rather thermal noise and interference contributed by the receiver. So when you calculate the second central moment, you're not calculating the second central moment of the delay function, but rather the second central moment of the delay spread function plus the noise. Because of that we decided to clip off the noise values and clip off the outliers and came up with a heuristic algorithm that calculated the second central moment [SLIDE 16]. And although the algorithm itself has no known statistical basis, it seems to lead to results which are consistent with the measurements. And that is all that one can expect from from any model. Thank you.

LINDSEY: Thank you Al. You know it's frustrating to prepare all of these viewgraphs and then not have enough time to go through them. I have to compliment the speakers on their diligent efforts and willingness to talk no longer than 20 minutes but I would like to spend a few minutes that we have to open it up for some discussions. If there are any questions and comments regarding any of the talks, I'd be happy to hear them. Is there a question from the back? [CUTOFF ...] It's fading ... we can hear you OK but maybe it's not being recorded. ... Do you want to use my mike?

JOHN TREICHLER: Now I've got to remember my question. A couple of things. One is, just for people in the audience who may in fact consider paying for experiments like Paul was talking about, my experience is the amount of time and energy taken to reduce the data is typically is 2, 3, 4, 5, sometimes 10 times more than it was to actually run the experiment and the experiment always costs 90% of the money and you end up with this tremendous pile of data and no time even to document it much less to reduce it. Another point though is, when you consider ... that is, when you actually get into designing the piece of equipment, you spend as much time with very high priced people arguing about what the data would have been had you been able to collect it, than you would going out to collect it in the first place. [LAUGHTER] Having been involved in a number of conversations over the

last 15 years, I heartily applaud efforts to go out and get the data to make it available.

The one point I did want to make of a technical nature is that I'm really amazed that you found no time variability in the last two ... I've been involved in a number of experiments over the last ten, fifteen years looking over microwave data, and say the 6 GHz band. We've found it very time-variable. It depended on time of day, the amount of wind that went through the medium. Now, most of this wasn't forest. It was like grassland and things like that, but it was very time-variable. First thing in the morning things were mirror-like and very steady, but just as soon as the winds came up you'd see time-variations of time scales in the order of fractions of a second and see Doppler spreads on the order of 10 Hz in almost no time. Now when we were first doing this work we went to a fellow named Yudlebee out in the mid-West who was doing back-scatter and forward-scatter and asked him why his data showed none of that time variation. He said "Ah! It wasn't necessary in our experiments, we just averaged it out." So you have to be a little careful when you look at the data because in fact it was there and they confirmed it but it was unimportant to figuring out how good soybeans were and so they didn't care.

SASS: Most of them were in the trunks in the UHF frequency band and in that regime you don't expect the wind to be blowing the trunks as Al pointed out. I did point out that in the trees in Connecticut we would have expected some in the leaves, I'm not sure why it's not there.

TREICHLER: In your results I'm just saying it's sort of surprising that given a factor of only 3 variance in frequency we saw a fairly significant effect.

LINDSEY: Oh yes, Phil ..

BELLO: I assume you took this data because there are some tactical communication links for example. But they wouldn't all be stationary. Now, can you use your data to predict Doppler spread due to motion because it'll be there in actual links. Did you take the impulse responses spaced close enough so you could piece them together?

SASS: That's exactly what we tried to do when we designed the system. We had a Doppler spread spec of 200 Hz, I believe it was and that was the ... we tried to make sure our sampling adhered to. So in fact we did a number of experiments where we did space TVIR experiments spatially.

BELLO: What about Doppler?

SCHNEIDER: The forest channel model doesn't presently incorporate the Doppler, although we have extended the model, at least heuristically, to look at radar returns from moving aircraft scanning over the top of canopy. However, it is fairly straightforward to put the Doppler shift into the multiple scattering model.

BELLO: But you haven't actually used your data to predict Doppler spread due to motion? Not yet?

SCHNEIDER: I guess I have to caveat that. When the delay spread functions were measured, in some cases they were repeated every four seconds, as Paul Sass said. These repeated measurements were then analyzed in a time series to determine the variation with time of the power in any of the delay bins. We found that - for nearly all of the data over the measurement intervals - which may extend over 15 or 20 minute intervals - based on stationary transmitters and receivers, there was no noticeable Doppler at all even though there may have been some slight wind up on top of the trees.

BELLO: I wasn't talking about that. I

was saying that in actual practice with one terminal moving, you will produce a Doppler spread. Have you ever used your fixed impulse response to reconstruct a time-variant impulse response due to motion? Have complex impulse responses been recorded?

SCHNEIDER: No.

BELLO: You haven't done that?

SASS: We have not, no.

LINDSEY: Presumably you'd have to move the transmitter relative to the receiver to get any Doppler spread measurement, would you not? Unless ....

SASS: They've taken them spatially at different positions.

LINDSEY: Oh, different positions.

SASS: That was the experiment we performed. We spatially ... that time-varying impulse response would be there if you were moving.

BELLO: In other words by taking impulse response measurements one foot apart you didn't find a complete, drastic change in impulse response?

SCHNEIDER: They have not yet been analyzed, Phil.

BELLO: Because I've noticed in many of these you don't have to move much before you get a gross change – the impulse response looks entirely different. You'd have to space your measurements very slightly apart in order to see this effect.

SCHNEIDER: We have made some preliminary calculations on the spacing, but that data has not yet been analyzed.

LINDSEY: Thank you. Any other comments? Yes, sir. That's Bob Peile. Could you give him a microphone.

ROBERT PEILE: With the benefit of hindsight and with today's technology, if you would set out with the object of finding which parameter is the dominant one in determining the effect on propagation, do you think you could have done anything with computer simulation better than measurement, i.e. if you started today?

SASS: In the forest channel we didn't have the model to do the simulation. You mean ....

PEILE: Supposing you started today and instead of a complete scientific characterization, if you started out with the question, which parameters are the dominant ones, clearly a first-order as opposed to second-order, my question is, could you have done anything but field measurements if you were starting today with today's technology of computer simulation ability?

SCHNEIDER: I think the problem is so multi-dimensional. There are some parameters that we anticipated ahead of time, but we don't even have to talk about hind sight because they had been already correctly anticipated. For example, the number of trees per acre, the tree trunk diameters, etc. As far as simulation is concerned, I don't know whether you are talking about perhaps simulating a forest with trees at certain positions and calculating the scattering from the trees and then adding them up on a Monte-Carlo basis.

PEILE: That's what I had in mind.

SCHNEIDER: You really shouldn't say that because what you're saying is that the volumes of scattering theory that the electromagnetics people have put together is better simulated. When you have random scatterers, hundreds and thousands literally of them, you really need to use sophisticated modeling techniques that are based on theoretical principles, rather than brute force techniques based upon Monte-Carlo approaches. Also, there are certain parameters that simply cannot be anticipated. For example, when we started, we didn't know whether rain might

have an effect on the attenuation through the canopy. We thought, "Well, when it rains, maybe we ought to see something." How would we have modeled that? For example, maybe we could have added a dielectric layer of rain water about the leaf or put a dielectric shell around the tree trunks or something like that. But as it turned out, there was no effect that was observed at all in that case. So I don't think in that situation we could have done it a priori. In that case, I think the measurement would have been the only way because of the complexity of the problem.

PEILE: So from that point of view, it was all worthwhile.

SCHNEIDER: Oh, yes, absolutely.

SEYMOUR STEIN: I have a question.

LINDSEY: O.K. Seymour Stein has a question. Seymour ....

STEIN: If you can catch your attenuation stuff, you've plotted it in dB per meter or dB per kilometer. That implies exponential attenuation. Is that what you observed with range?

SCHNEIDER: dB per meter.

STEIN: What kind of ranges did you test over?

SCHNEIDER: We tested over 3 ranges – 3 distance measurements. The path loss measurements were made over 3 parallel paths, in essentially the same area of the forest in order to maximize the homogeneity over the paths. The shortest path was about 300 ft, the other one around 500 ft, and the third about 1000 ft. After we measured the power, we looked at the profile – the tree trunk number density along the path – and normalized that out in order to calculate the specific attenuation. Only three paths were measured and they were differential measurements to get the specific attenuation.

STEIN: Are you talking about your simulation or the actual measurements.

SCHNEIDER: I'm talking about the actual measurements.

STEIN: OK. Your much earlier conclusion about invalidating the up, over and down model, I thought that that model applied to how you propagate over longer ranges in the jungle, not the short ranges.

SCHNEIDER: Well, the up-over-and-down model is a consequence of the coherent component. It arises as a consequence of being able to replace the random scatterers within the forest by some equivalent or effective dielectric slab and then looking for the critical angle of internal reflection. That's the angle at which the lateral wave propagates over the tree tops. As you move up in frequency over 200 MHz, the size of the scatterers and the separation between them become such that the coherence needed to launch a surface wave on a interface of two disparate dielectric constants is lost. You are not able to maintain the phase coherence from the point at which the wave is launched. That's the theoretical reason why the lateral wave would probably not ever be observed at frequencies above 300 MHz. In spite of the measurements we took – perhaps we didn't go out long enough, but we went out as far as we possibly could within a homogeneous forest – we did not observe any.

STEIN: What would be your conclusion about transmitting from a terminal down on the ground to a terminal up in the air some distance away? Are you saying that there's no reasonable way to get from the top down without a horrible attenuation?

SCHNEIDER: When you say communicating, do you mean a path between two antennas at the same height, or disparate height, one inside and one outside?

STEIN: One inside and one outside.

SCHNEIDER: No, I don't think there's any low-loss mechanism that's going to be helpful here. It's just simply going to be specific attenuation at that angle through the forest. It will depend on angle.

STEIN: At these frequencies?

SCHNEIDER: Yes, at these frequencies, Seymour.

LINDSEY: Any further comments or questions? Yes, Sir.

SCHNEIDER: I had two questions for Phil. When you spoke about the tapped delay line representation for the HF channel model, you mentioned that the spacing of the taps had to be essentially the reciprocal of the bandwidth and, because of the total delay spread of the channel, that it would take perhaps thousands of taps to represent that channel. The question I have is, when you've looked at all these alternative channel models in your '63 paper that were effectively equivalent representations - tapped delay line models and different transfer functions, power series, expansions and so forth - are you led to believe that perhaps for the HF channels where only a few of the taps are likely to be active in the vicinity of the modes, that perhaps a different representation might afford a more tractable channel model representation?

BELLO: Well, it turns out that not just the few number are active. The individual propagation modes due to dispersion for example, even on an undisturbed channel at short ranges can be several hundred microseconds. But it's not small.

SCHNEIDER: Well, the differential delay - let's say between a 1-hop or a 2-hop path or between a high-angle ray or a low-angle ray - can typically be on the order of milliseconds - that differential delay. So in each mode you may have a few but there may be a gap.

BELLO: Yes, I went through the thing too quickly but each propagation mode would be modeled by a tapped delay line which would have hundreds of tap. The modes would be separated by delays of the order of milliseconds.

SCHNEIDER: The second question I had - you mentioned that one of the difficulties especially with an HF channel is the fact that the external noise tends to be non-gaussian.

BELLO: Non-stationary, and non-Gaussian too.

SCHNEIDER: I would suspect that because it's non-gaussian, that would be an advantage to the communicator as opposed to a disadvantage.

BELLO: Well, just think of what happens *if you measure the noise power* over short intervals of times, it fluctuates substantially. So, suppose you're trying to acquire a received spread-spectrum signal, you have to set your thresholds higher to fix your false-alarm rates. It's clearly a loss to have this non-stationary fluctuation.

SCHNEIDER: I wasn't emphasizing non-stationarity. I was talking about the non-gaussian aspect of it.

BELLO: Well, I wasn't talking about non-Gaussian, I was talking about non-stationarity.

SCHNEIDER: Do you think the fact that it is non-gaussian is an advantage?

BELLO: I haven't thought about the Gaussian or non-Gaussian character. In many cases you're averaging so long that the short term statistics are nearly Gaussian. It's just that the power varies rapidly over short intervals of time.

LINDSEY: I guess my comment with regard to non-Gaussian statistics, I think it

does help the communicator. That's an opinion that I feel intuitively and I think one could quantify that if used properly.

SCHNEIDER: I agree with you.

LINDSEY: On the other hand you open up the issue of theorists who like to deal with complex Gaussian envelopes you know. And what are you going to do if you don't have that and that's the issue. I think the statistical model which I talked about certainly indicated that statistics aren't Gaussian, and even in some of the measurements that Phil showed. Rayleigh is not the correct model. But it fits in some areas and it's a good analytically tractable approach to playing games to get numbers to design a modem, or whatever. We may be finished in fact .... Oh, no, we have another comment.

STEVE STEARNS: This is Steve Stearns. I have a question for Phil Bello, a different aspect of HF propagation. The scattering function formulation is useful for channel models where you have a single input and a single output, but another application, radio direction where you are using an array of antennas and you're interested in spatial diversity as opposed to mode diversity at a single receiving antenna, are there any appropriate extensions to the model that you can use for this case? Let me add a reason why I'm asking this question. The ionosphere is not modeled by a spherical layer at uniform height that is centered on the axis or center of the earth. Rather it has tilts in it and consequently the angle of arrival that's measured does not correspond at all to the great circle direction, or line varying to a source, and moreover it fluctuates with time. I'm asking this question with regard to radio direction finding systems at HF.

BELLO: I didn't have time to go into the full scale modeling, but I think Bill Lindsey was indicating that the scattering function can be generalized. You can have a scattering function in delay, Doppler and direction cosine, or angle. And really that's what Bill was talking about.

LINDSEY: I think it's exactly what my model was saying that you have to take into account the direction, not only just one direction that energy is coming at you but you need to take into account direction that energy is coming from the other two. Namely, we've probably looked at two dimensions, namely delay, although to me that's the direction of propagation of foresight of one antenna to another, and then if you have to worry about the other two coordinate axes which you should I think, then you'll begin to see other things that may or may not be important. It may not be, but I think it is.

The ionosphere, I believe radio physicists are still playing with what to do with that problem but I think we are beginning to learn quite a bit of it. And the effects of the layers can be handled in a transport theoretical formulation which is the direction that I had taken to worry about multiple reflections between the various boundaries associated with the ionosphere.

BELLO: Let me point out that the propagation physicists use something called a mutual coherence function and talk about basically the correlation of the fields separated in space and time for different transmitted frequencies. The three-dimensional Fourier transform of this correlation function yields the three-dimensional scattering function. The work by Nikish I mentioned in my viewgraphs (in fact there will be a sheet of references on that), derives this three-dimensional spatial-time-frequency correlation functions. There are other references in it, too.

LINDSEY: The Soviets are playing with that concept too. Soroski, and take-offs on Kolmorogof's work, and so on. Seymour is looking like he'd like to get involved in something here.

STEIN: With respect to direction finding you'd find that going back to this literature, there's a lot of work on angular ... and we are not talking about mutual coherent function. We are talking about multiple modes which are quite discrete. And it's really the angular red that's the problem and nobody knows where the modes are coming from in any way that will enable you to do any theoretical predictions. That's the problem. I have one final question that I'd like to take back to the measurement problem.

LINDSEY: O.K. Seymour ....

STEIN: One of the practical issues when you get to very wide-band direct sequence is the question of what you have to do to build a ray combiner. Now how many taps do you need. Can you build a spark combiner. Were any of the measurements done, (you had multiple resolution) so that you had what amount to simultaneous data on the same instantaneous response with different resolutions to tell whether they looked like discrete paths, or whether the paths kept breaking up if you went to finer and finer resolutions?

SASS: We did some measurements like that but again they were in the forested channel. The system right now is being used by the Department of Commerce to make some measurements in the Denver area. So I think that's the environment you'd want that data in, in an urban environment.

STEIN: What about interpretations for the forest channels?

SCHNEIDER: To the resolution of the equipment, we never noticed any discrete scatterers.

STEIN: Well your pictures showed what looked like discrete scatterers.

SCHNEIDER: No, that's because that's the chip resolution. But there was no coherence or correlation between adjacent lines ever. You couldn't identify any discrete scatterers.

STEIN: Now let's not talk about discrete scatterers, but discrete delays however they are coming about. Did it keep breaking up as we went to final resolution? If you did the same impulse response with two different resolutions, very close together in time, did you find breakup, or did you find ....

SCHNEIDER: You mean a randomness associated with the delays on adjacent bins? Yes, they were always completely randomized.

STEIN: Interesting. Thank you.

BELLO: I had one more thing to add to what Seymour said. My comments apply to the disturbed channel. You get to the disturbed channel, that's where the scattering function means something. The undisturbed is what Seymour was talking about.

LINDSEY: O.K. with that I'd like to conclude this session. On behalf of the Communication Sciences Institute I'd like to thank the speakers Paul Sass, Al Schneider, Phil Bello, and Ken Wilson for taking time out of their regular busy schedule to prepare viewgraphs and present the materials, and come here and share with us the knowledge that they've gained. On the other hand it was my pleasure to be involved with this session and thank the participants in the audience in dealing with this session. Finally I'd like to hand it over to our leader, Bob Scholtz, and see if he has any direction to pass out.

SCHOLTZ: I don't think I really have any comments except you can see there are some interesting discussions that will always de-

velop at the end of sessions. So the remaining panelists, I've been noticing the talks creep. You know, the first one is 20 minutes, the next 25, the next 30. It happens every session and in the next session everyone adds another 5 on. So what I'd like to do is try to cut down the presentations. The discussions will come and you'll get to say everything you wanted to say but it will be a slightly different environment of give and take. So let's work on that for the next two sessions. Thank you very much. Cocktails and entertainment are waiting, hors d'oeuvres and all sorts of things, so let's get going.

U.S. ARMY COMMUNICATIONS-ELECTRONICS COMMAND
WIDEBAND PROPAGATION MEASUREMENT SYSTEM

(Designed and Built By SRI International)

CECOM Project Leader:
Kevin Lackey (201) 544-5680

SRI Project Leader:
Boyd Fair (415) 859-3136

SYSTEM CAPABILITIES FACT SHEET

Overview

The Wideband Propagation Measurement System (WPMS) is specially designed and built to measure the very wide-bandwidth radio propagation characteristics of communications channels that interest the U.S. Army and other U.S. government agencies. These channels primarily consist of, but are not limited to, forested and urban environments. The goal of the WPMS program is to develop and verify a communications-propagation model for spread-spectrum signals that are propagated through these channels.

The WPMS consists of two mobile vans, one containing the transmitter subsystem and one containing the receiver and data-acquisition subsystem. The WPMS is designed to transmit and receive two simultaneous spread-spectrum signals that can be centered at any frequency between 200 MHz and 2000 MHz. These spread-spectrum signals are generated by modulating the rf carrier with a pseudo-random noise (PN) code which is then radiated and received by the receiver system.

The received signal is cross correlated with a time-delayed replica of the transmitted waveform which generates an amplitude-vs-time waveform called a *time-varying impulse response* (TVIR) of the communications channel. The TVIR is a representation of the response of the communications channel to an impulse input signal, perturbed by any scattering, multipath, or absorption mechanisms.

Although the WPMS is an asynchronous system, the TVIR data also provides a relative time-of-arrival measurement of the multiple signal paths (multipath) generated within the channel. The TVIR data are calibrated in absolute received power.

The resulting digitized spread-spectrum data can then be processed to measure several facets of the received signal: the level (useful for computing path loss), Doppler shift, delay spread (a measure of the amount of existing multipath, which is useful in computing the

maximum data rate for a communications channel), and the power spectrum of the transmitted signal at the receiver. These measurements can be made at any antenna height, from man-pack level up to 65 feet with either directional (vertical, horizontal, or circular polarization) or omni-directional (vertical polarization) antennas. Mobile operation is also possible using roof-top antennas.

The WPMS system is semi-automatic. Previously generated experiment files are used to automatically control the setup of the transmitter and receiver hardware and to gather the data and record it onto magnetic tape. The raw data is then processed off-line to provide a set of calibrated summary data that can be further analyzed by scientists and engineers interested in specific aspects of the spread-spectrum radio propagation problem.

The specific capabilities of the WPMS are summarized in the following table.

SYSTEM CAPABILITIES

GENERAL	
Carrier frequency range	200 to 2000 MHz
Delay-spread range	-1 to 20 us
Delay-spread resolution	2 ns
Doppler-spread range	-15 to 240 Hz
Doppler-spread resolution	<2 Hz
TVIR amplitude resolution	0.1 dB
TVIR multipath amplitude resolution	-20 dB min
Measureable path loss	155 dB min
ANTENNAS	
Omnidirectional	Biconical
Directional	Crossed LPAs
Transmitter polarization	V, H, RCP
Receiver polarization	V, H, RCP, LCP
Azimuth	0 to 360°
Height	5 to 65 ft
TRANSMITTERS	
Number	2
Frequency range	
Transmitter No. 1	200 to 1050 MHz
Transmitter No. 2	700 to 2000 MHz
Power output	100 W max
Modulation	PN sequence Bi-phase modulation or CW
Clock rate	50, 125, or 250 MHz
Instantaneous null-to-null bandwidth	100, 250, or 500 MHz
PN code length (chips)	255, 511, 1023, or 2047
Control	Local or receiver computer
RECEIVER	
Number	2
Frequency range (each channel)	200 to 2000 MHz
Sensitivity (23 dB SNR in 50 kHz BW)	-95 dBm
Modulation	PN sequence Bi-phase modulation or CW
Clock rate	50, 125, or 250 MHz
Instantaneous null-to-null bandwidth	100, 250, or 500 MHz
Code length (chips)	255, 511, 1023, or 2047
Control	Local computer
Instantaneous dynamic range	50 dB
Input signal range (at van wall)	0 to -95 dBm
AGC (receiver computer controlled)	Stepped in 1-dB increments

U.S. ARMY COMMUNICATIONS-ELECTRONICS COMMAND WIDEBAND PROPAGATION MEASUREMENT SYSTEM

(Designed and Built By SRI International)

CECOM Project Leader:
Kevin Lackey (201) 544-5680

SRI Project Leader:
Boyd Fair (415) 859-3136

EXECUTIVE SUMMARY

Introduction

The Department of Defense (DoD) is interested in increasing the bandwidth of transmitted signals for military applications for three principal reasons:

- To reduce vulnerability to jamming
- To increase data transmission rates
- To reduce the probability of interception.

To achieve these goals, systems that use frequency hopping and direct-sequence (pseudonoise) spreading are becoming more common in a variety of military, as well as civilian, applications. The military is developing or experimenting with systems such as JTIDS, PLRS, SINCGARS, and Packet Radio. Spread spectrum techniques are also employed by NASA's TDRSS and DoD's NAVSTAR GPS satellites.

For all these systems, instantaneous signal bandwidths can reach 10's of MHz. The usefulness of such wide bandwidth systems in ground environments is uncertain in many cases because they must be deployed in complex channels—such as forests and urban environments. For each of these channels, a number of parameters (including delay spread, Doppler spread, and coherence bandwidth) determine the channel's ability to support broadband communication without degrading the system's performance.

Channel-induced degradation is a particularly important problem to the Army. The Army must communicate while *on-the-move*, over all types of terrain covered with every conceivable type and density of vegetation. Previous wide-bandwidth measurements of forested communication channels have shown that the medium may be characterized by highly structured multipath. This multipath, combined with channel fading resulting from vehicular motion, can significantly affect the performance of existing communication systems and the design of future ones.

To determine the effect of this communication environment on spread-spectrum systems, three approaches are possible:

- Evaluate system performance based on theoretical models of the channel and the communication system
- Evaluate system performance based on empirical (statistical) models of the channel (verified by experimental measurements on actual channels) and a theoretical model of the channel
- Evaluate actual systems in the actual environments of interest.

All of these approaches have their use in developing and testing communication systems. The WPMS is a system that can obtain data necessary for the second and third approaches.

Background

Channel Measurement Philosophy

The channels of interest can be modeled as randomly time-varying systems, for which the only meaningful description is statistical. The statistics of a given characterization, however, may be deterministic, depending on unknown external parameters, such as season, moisture content, forest density, and the like.

The ultimate rationale behind a channel characterization program is to evaluate the performance of a communication system over channels similar to those likely to be encountered in practice. This evaluation requires a characterization that encompasses "typical" as well as extreme channels. However, because Army communication systems must operate over widely varying conditions, measuring all the possible generic types of channels is impractical. Therefore, understanding the underlying physical phenomena of the channels is necessary, so that the results of the measurement program can be extrapolated to any channel of interest.

This extrapolation is easier if the effects of each of the various external parameters can be separately identified. Then, a relationship between the statistical channel characterization and the external parameters can be established. If identifying an effect with a particular external parameter is not possible, it is imperative, at least, to try to control as many variables as possible. Such control is best accomplished by using a simple channel such as a "standard" forest without any complicating factors such as non-homogeneity, variable tree sizes or types, hilly terrain, or similar conditions. These factors can then be added later by extrapolating the statistical model developed from the basic set of measurements. This approach is preferred over measuring a number of complex channels, because identifying cause with effect is difficult when too many variables are changing at once.

Time-Varying Impulse Response

The purpose of the WPMS system is to measure the Time Varying-Impulse Response (TVIR) or an equivalent functional representation, known as the Output Delay-Spread Function (ODSF) of the channel. The ODSF and TVIR are entirely equivalent; the only difference is the way in which the time and delay variables are referenced to the time the impulse is applied to the communications channel. The delay variable of the ODSF is referenced (equals zero) to the time the impulse is applied to the communications channel. Therefore, the ODSF is equivalent to an oscilloscope display of the response of the system to a unit impulse for which the time origin of the oscilloscope is set to the time the impulse was applied to the channel. This input time is the only place "absolute" time appears in the ODSF.

The function measured by the WPMS is actually a time-shifted version of the ODSF — offset in time by an unknown amount because of the unsynchronized nature of the WPMS measurement system.

System Overview

The WPMS uses a sliding-correlator architecture (SCA) to measure the TVIR of a propagation channel. In this approach, a wide-bandwidth pseudonoise code is used to modulate a UHF carrier at the transmitter. This composite signal is then cross-correlated at the receiver with a duplicate code. In the receiver, however, the clock for the shift register generating the duplicate code is stepped (slid) one chip per PN code period relative to the transmitter clock; thus, if we compared the transmitter and receiver codes as a function of time, assuming they are being continuously repeated, we would observe the codes stepping (sliding) past each other.

The correlators are capable of operating with PN-code lengths of 255, 511, 1023, and 2047 chips at code speeds of 50, 125, and 250 Mbits/sec. To shorten the correlation times and to reduce the speed at which the Analog/Digital connectors (A/Ds) must take data, four parallel correlators are used in each receiver channel. The second, third, and fourth correlators correlate the received waveform with a copy of the PN-code that has been shifted relative to the PN-code in the first correlator by 1/4, 1/2, or 3/4 chips, respectively.

This approach for system implementation enables the basic design objectives for the WPMS to be met at a reasonable cost. Selecting an SCA, however, is not without its problems. The use of the SCA must be carefully controlled by the DDAS to ensure that valid data are collected regarding the time variability of the channel. In fact, all the requirements cannot be satisfied simultaneously (at reasonable cost). As a compromise solution to this problem, a multi-mode system of operation [controlled by the Digital Data Acquisition System DDAS)] has been developed that permits the experimenter a tradeoff between conflicting requirements in:

- TVIR resolution in the time domain
- TVIR bandwidth
- Duration (multipath-delay window) of the TVIR in the time domain
- Total path loss measurable by the system.

The WPMS is functionally divided into two subsystems. The first part, the Receiver/Digital-Data-Acquisition-System (RX/DDAS), contains all the hardware necessary to receive, sample, process, and record the various channel-probe signals that are used to characterize time-varying radio propagation channels. The second part of the WPMS consists of the transmitter (TX) and its microcomputer controller.

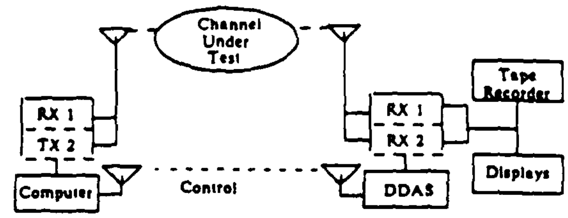


Figure 1. WPMS Block Diagram

The RX/DDAS has been designed to provide automated control of the experiment and a "friendly" user interface for conducting and designing channel-probe experiments. The application-level software for this interface is designed to verify at each menu-driven prompt that the user enters "reasonable" information or specifies "reasonable" parameters for the experiment. These software checks help minimize wasted experiment time caused by user errors and minimize the probability of recording channel-probing data that are invalid or not meaningful.

The TX subsystem normally acts as a *slave* to the RX/DDAS. That is, the transmitters are controlled by the transmitter van controller (microcomputer), which receives its commands from the RX/DDAS. A VHF communications system is used between the two subsystems to transfer commands, acknowledgements, and status information. When an experiment is executed at the RX/DDAS, the appropriate commands are sent to the TX controller. This controller locally controls the transmitter hardware consistent with the RX/DDAS commands. After the transmitter has been set up and the data recorded by the RX/DDAS, the TX controller sends the TX status information to the RX/DDAS for inclusion in the header block of the recorded data.

Conclusion

In this executive summary we have presented a high-level overview of the WPMS, built by SRI under contract to the US Army for the purpose of gathering spread-spectrum propagation data to characterize forested and urban communication channels. This characterization will be accomplished by using very-wide-bandwidth, probe signals at frequencies ranging from 200 to 2000 MHz. A summary of the systems capabilities is shown on the tables that follow.

These capabilities meet or exceed all of CECOM's requirements for the WPMS. This state-of-the-art system provides the ability to measure ground-to-ground communication channel characteristics with very high resolution, using very-wide-bandwidth probe signals. These characteristics will be used to verify and improve computer models that will, in turn, be used to evaluate the performance of the Army's communication systems of the future.

BANDWIDTH/CODE-LENGTH PARAMETERS

Bandwidth (MHz)	Chip Length (ns)	Code Length (Chips)	Code Length (ms)	Time to Acquire One TVIR (ms)	Number of TVIRS to Integrate	Time to Acquire Complete TVIR (ms)	Number of TVIRs in 128k Samples	Doppler Bandwidth (Hz)
50	20	255	5.100	1.306	2	2.611	512	191.483
50	20	511	10.220	5.233	1	5.233	256	95.554
50	20	1023	20.450	20.951	1	20.951	128	23.865
125	8	255	2.040	2.089	4	10.522	512	239.354
125	8	511	4.088	2.093	1	2.093	256	238.885
125	8	1023	8.184	8.390	1	8.380	128	59.663
125	8	2047	16.376	33.538	1	33.538	64	14.908
250	4	255	1.020	0.261	8	2.089	512	239.354
250	4	511	2.044	1.047	2	2.093	256	238.885
250	4	1023	4.092	4.190	1	4.190	128	119.326
250	4	2047	8.188	16.769	1	16.769	64	29.817
CW	n/a	n/a	n/a	n/a	n/a	n/a	n/a	250.000

System Capabilities

GENERAL	
Carrier frequency range	200 to 2000 MHz
Delay-spread range	-1 to 20 us
Delay-spread resolution	2 ns
Doppler-spread range	-15 to 240 Hz
Doppler-spread resolution	<2 Hz
TVIR amplitude resolution	0.1 dB
TVIR multipath amplitude resolution	-20 dB min
Measureable path loss	155 dB min
ANTENNAS	
Omnidirectional	Biconical
Directional	Crossed LPAs
Transmitter polarization	V, H, RCP
Receiver polarization	V, H, RCP, LCP
Azimuth	0 to 360°
Height	5 to 65 ft
TRANSMITTERS	
Number	2
Frequency range	
Transmitter No. 1	200 to 1050 MHz
Transmitter No. 2	700 to 2000 MHz
Power output	100 W max
Modulation	PN sequence Bi-phase modulation or CW
Clock rate	50, 125, or 250 MHz
Instantaneous null-to-null bandwidth	100, 250, or 500 MHz
PN code length (chips)	255, 511, 1023, or 2047
Control	Local or receiver computer
RECEIVER	
Number	2
Frequency range (each channel)	200 to 2000 MHz
Sensitivity (23 dB SNR in 50 kHz BW)	-95 dBm
Modulation	PN sequence Bi-phase modulation or CW
Clock rate	50, 125, or 250 MHz
Instantaneous null-to-null bandwidth	100, 250, or 500 MHz
Code length (chips)	255, 511, 1023, or 2047
Control	Local computer
Instantaneous dynamic range	50 dB
Input signal range (at van wall)	0 to -95 dBm
AGC (receiver computer controlled)	Stepped in 1-dB increments

U.S. ARMY COMMUNICATIONS-ELECTRONICS COMMAND WIDEBAND PROPAGATION MEASUREMENT SYSTEM

(Designed and Built By SRI International)

CECOM Project Leader:
Kevin Lackey (201) 544-5680

SRI Project Leader:
Boyd Fair (415) 859-3136

SAMPLE DATA

Background

The data contained herein are actual data taken in a forest at Ft. Lewis, near Seattle, WA. The WPMS was used to collect and record a very-large volume of wide-bandwidth and CW propagation data on many different paths within the forest. Measurements were made at antenna heights ranging from 10 ft to 65 feet over path lengths from 270 ft to 6000 ft. The majority of the data were gathered for both vertical and horizontal polarizations, however some crossed-polarized and circularly polarization data was also taken. The radio frequencies (rf) used were 300 MHz, 400 MHz, 850 MHz and 1050 MHz. These carriers were modulated to produce null-to-null bandwidths of 500 MHz for most of the measurements.

The figures that follow are representative samples of the data available form the WPMS. The photographs are near-real-time data displays available in the field during the data-gathering portion of the experiment. The graphs that are included are only some of the ways the *processed data* can be displayed using the post-processing programs developed for the WPMS.

1. Figure 1. Real-time displays of Time Varying Impulse Response (TVIR) and Delay Spectrum for two received signals. Note Wideband Received Signal Level (WRSL), Narrowband Received Signal Level (NRSL), and TVIR Received Signal Level (TRSL), and calculated Path Loss displays. Discrete frequencies in the spectrum displays are the result of rf carrier leakage and PN-code unbalance in the transmitter and receiver.

2. Figure 2. A real-time display of one of the TVIR displays. The trace on top shows the entire multipath delay-spread range for the code-length chosen for the experiment. The bottom display is an expanded version of the top trace centered at the tick mark shown (movable) on the top display.

3. Figure 3. A real-time display of one of the Delay Spectrums. Other types of displays can be presented on this device. These include the relative phase angle-vs-time of the received signal, the individual In-phase and Quadrature components of the TVIR, a polar display of TVIR amplitude and a spectrum display of the Doppler shift of the received signal.

4. Figure 4. A post-processed display of expanded TVIRs for vertical and horizontal polarization for a free-space path over a logging road within Ft. Lewis.

5. Figure 5. A post-processed display showing the polarization dependence of spread-spectrum signals when propagated through a trunk-dominated forest at Ft. Lewis.

6. Figure 6. A post-processed display showing the polarization effect on delay-spread encountered on many paths within the forest at Ft. Lewis.

7. Figure 7. A post-processed display of the calculated excess path loss encountered within the Ft. Lewis forest at vertical and horizontal polarization as a function of frequency and path length for spread-spectrum signals.

8. Figure 8. A post-processed display of the calculated excess path loss encountered within the Ft. Lewis forest at vertical and horizontal polarization as a function of frequency and path length for cw signals.

9. Figure 9. A forester characterizing the Ft. Lewis forest. These characterizations are critical to the evaluation of computer models of radio wave propagation in forests.

10. Figure 10. The WPMS receiver/data acquisition van gathering data at Ft. Lewis, WA.

11. Figure 11. The WPMS transmitter van radiating two simultaneous spread-spectrum signals at Ft. Lewis. Note the automatically raised and lowered antenna/tower subsystem.

12. Figure 12. The Digital Data Acquisition System used to process and record the experimental data and to control the WPMS transmitters, receivers and antennas.

13. Figure 13. The WPMS transmitter equipment mounted inside the transmitter vehicle.

14. Figure 14. The WPMS receiver systems prior to installation in the receiver van.

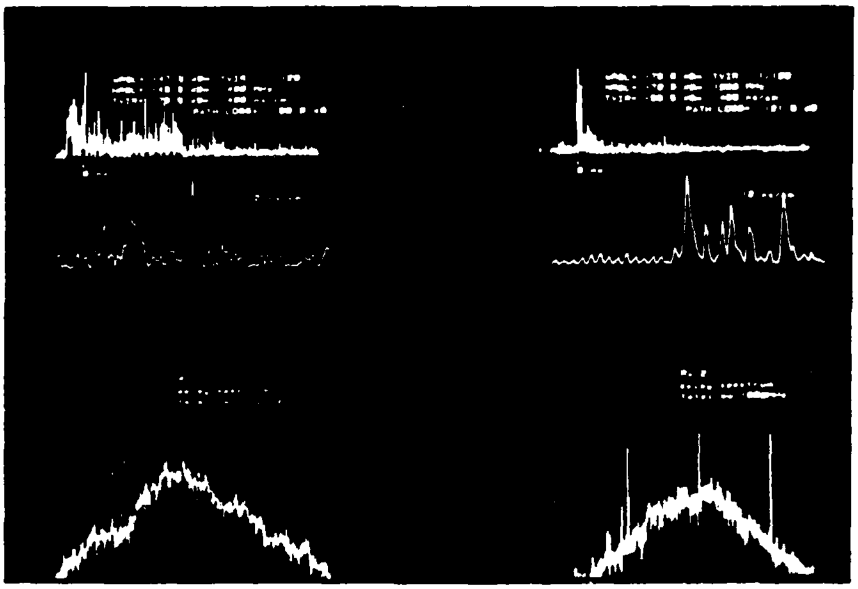


Figure 1.

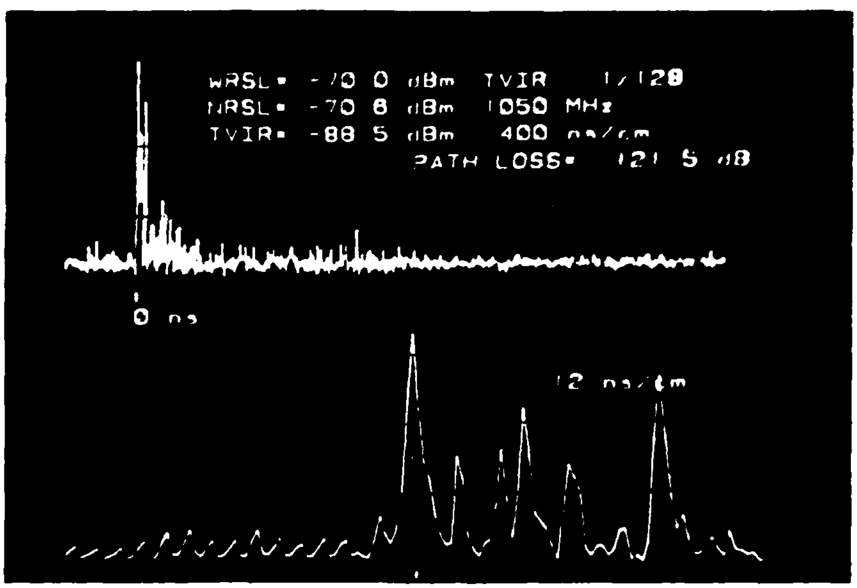


Figure 2.

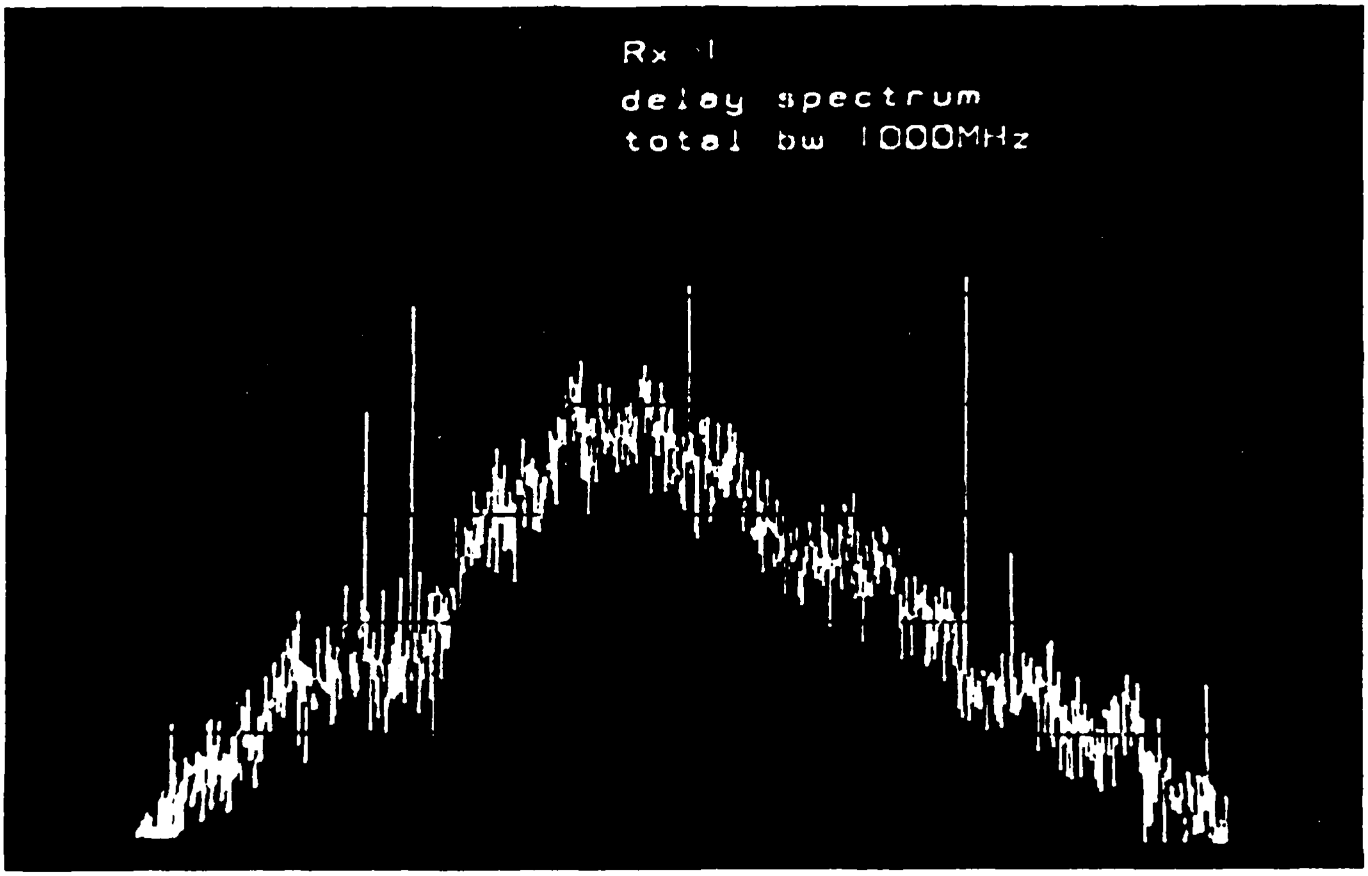


Figure 3.

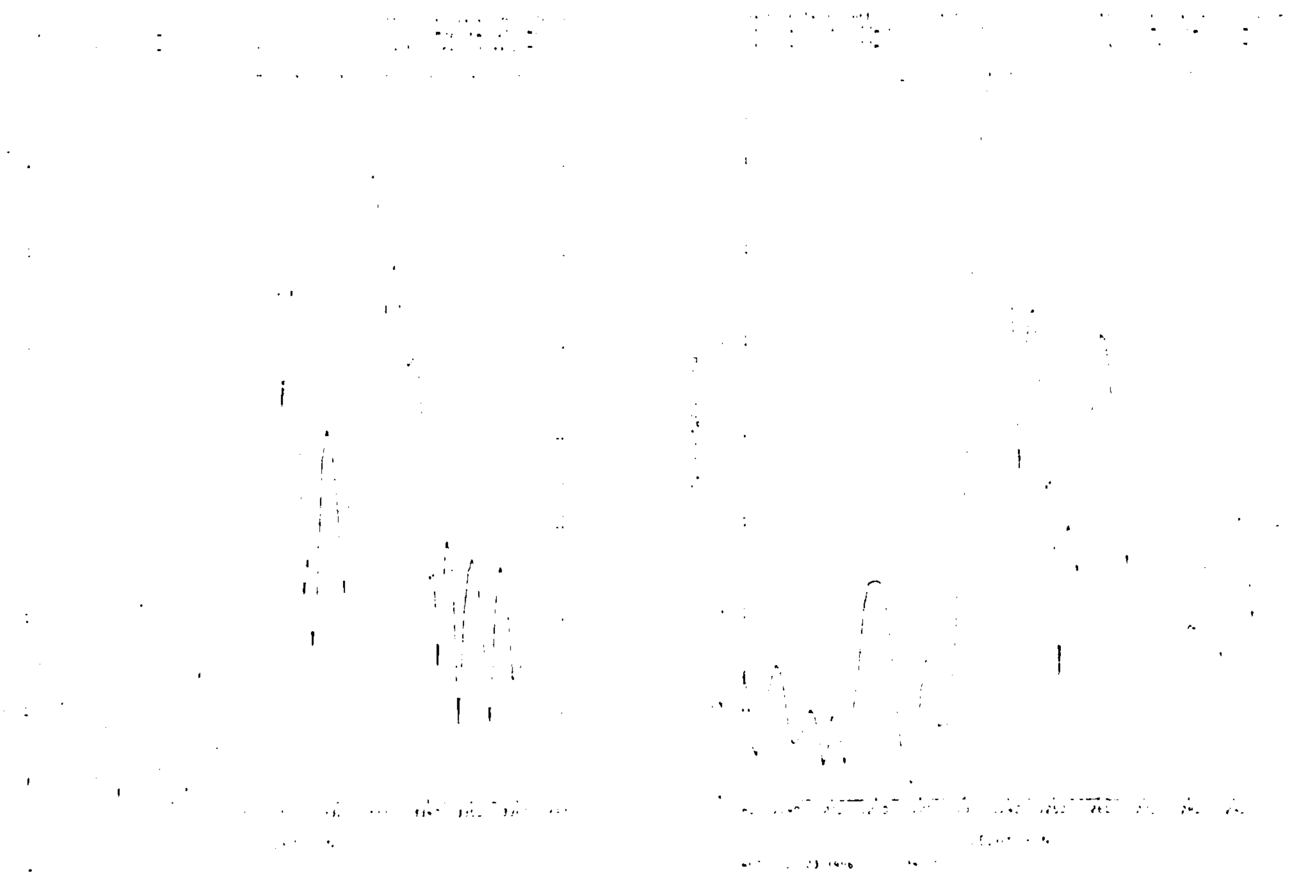


Figure 4.

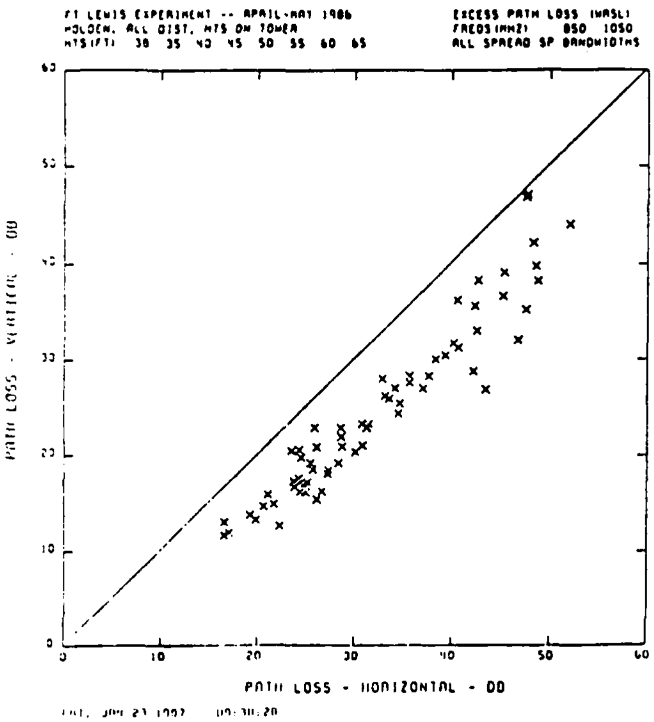


Figure 5.

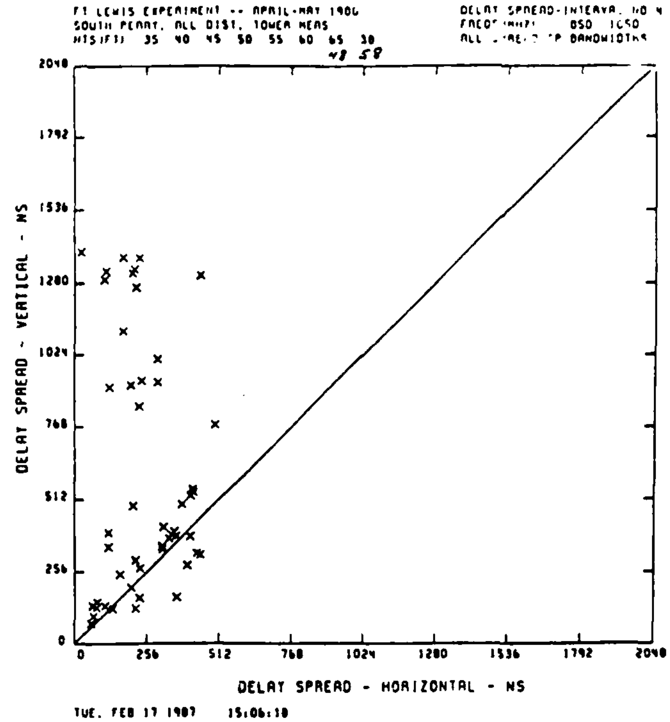


Figure 6.

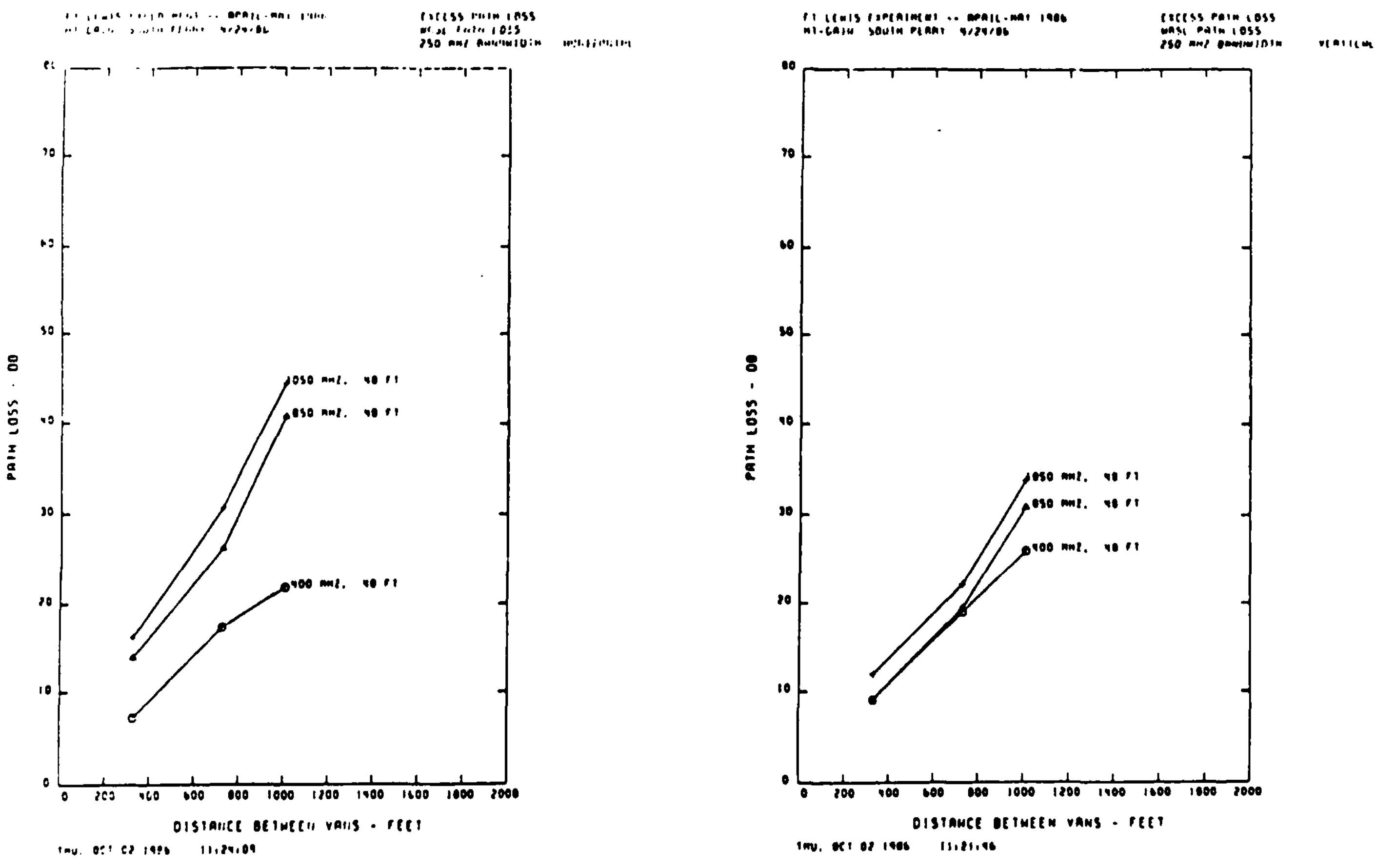


Figure 7.

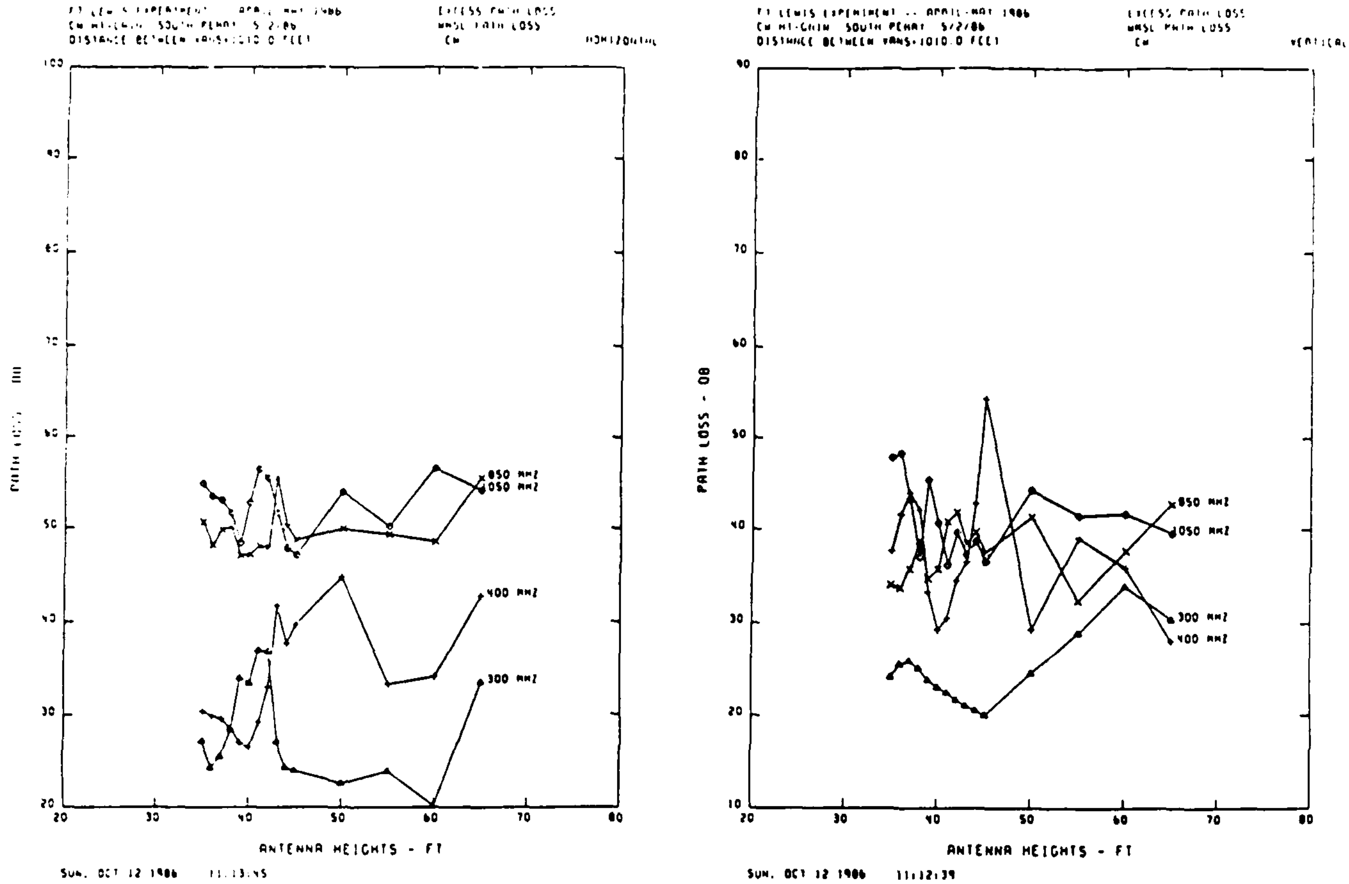


Figure 8.

Figure 9.

Figure 10.

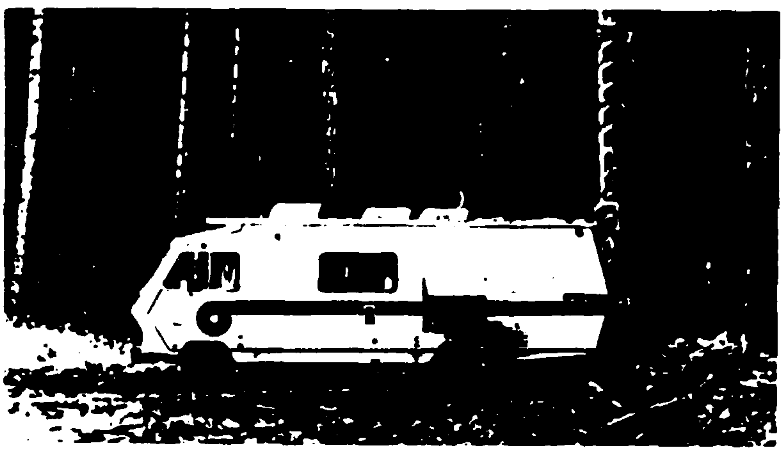


Figure 11.

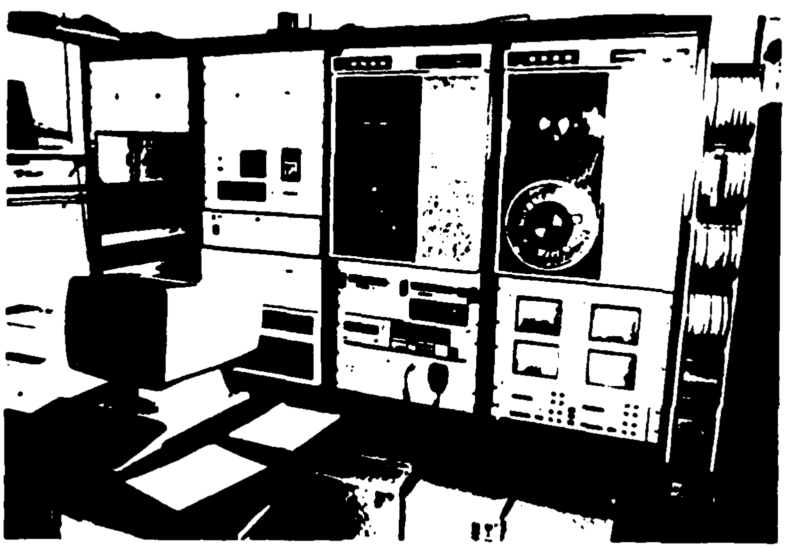


Figure 12.

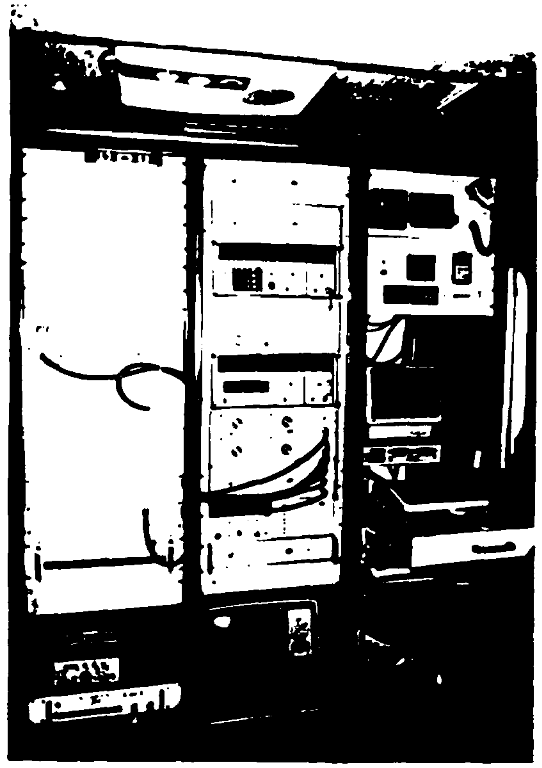


Figure 13.

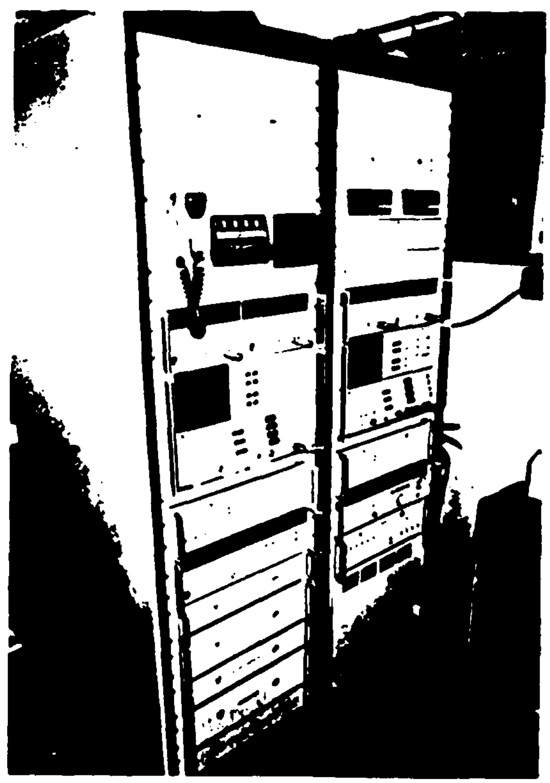


Figure 14.

CECOM Project Leader:
Kevin Lackey (201) 544-5680

SRI Project Leader:
Boyd Fair (415) 859-3136

BIBLIOGRAPHY

Antenna Engineering Handbook, Jasik, H., ed. (McGraw-Hill Book Company, New York, New York, 1961)

Bentley, P. B., "Frequency Dispersion of a Log-Periodic Antenna, "Technical Note, Contract DAAB07-82-C-J097, SRI Project 4520, SRI International, Menlo Park, CA (February 1983)

Brown, G. S. and W. J. Curry, "A Theory and Model for Wave Propagation Through Foliage," *Radio Science*, Vol 17, No. 5, pp. 1027-1036, (September 1982)

Dence, D. and T. Tamir, "Radio Loss of Lateral Waves in Forest Environments," *Radio Science*, Vol. 4, No. 4, pp. 307-318 (April 1969)

Fair, B. C. and M. S. Frankel, "Design Plan for Wideband Propagation Measurement Data/Acquisiton and Recording System," SRI Design Report 1, Contract DAAB07-82-C-J097, (September 1982)

Fair, B. C., R. A. Nelson and M. S. Frankel, "Wideband Propagation Measurement/Data Acquisition and Recording System," *Semiannual Status Report* 1, Contract DAAB07-82-C-J097, SRI Project 4520, SRI International, Menlo Park, CA (August 1983)

Fair, B. C., "Wideband Propagation Measurement/Data Acquisition and Recording System," Semiannual Report 2, Contract DAAB07-82-C-J097, SRI Project 4520, SRI International, Menlo Park, CA (May 1985)

Frankel, M. S., "L-Band Forest Experiments," Packet Radio Temporary Note 254, SRI International (May 1978)

Head, H . T., "The Influence of Trees on Television Field Strengths at Ultra-High Frequencies," *Proc. IRE*, Vol. 48, No. 6, pp. 1016-1029 (June 1960)

Herbstreit J. W. and W. Q. Crichlow, "Measurement of the Attenuation of Radio Signals by Jungles," *J. Research NBS Section D (Radio Sci.)*, Vol. 68D, No. 8, pp. 903-906 (August 1964)

Hodges, J. C., G. A. DesBrisay, C. L. Rino, B. C. Fair and J. Owen, "Site Selection Activities for the WPMS," Semiannual Report 1, Contract DAAB07-85-C-K569, SRI Project 8899, SRI International, Menlo Park, CA (November 1986)

Hufford, G. R., L. Hubbard, J. Pratt, S. Adams, Paulson and P. Sass, "Wideband Propagation Measurements in the Presence of Forests," CECOM Technical Report 82-CS029-F, (January 1982)

Hufford, G. A., Longley and W. Kissick, "A Guide to the Use of the ITS Irregular Terrain Model in the Area Prediction Mode," NTIA Report 82-100, US Dept. Of Commerce, (April 1982)

Josephson B. and A. Blomquist, "The Influence of Moisture in the Ground, Temperature and Terrain on Ground Wave Propagation in the VHF-Vand," *IRE Trans. Antennas and Propagation*, Vol. AP-6, No. 2, pp. 169-172 (April 1958)

"JTIDS Ground Foliage Propagation Tests," Final Report, Contract F19628-75-C-0205, FR80-16-503, Hughes Aircraft Company, Fullerton, CA, (April 1980)

LaGrone A. H. and C. W. Chapman, "Some Propagation Characteristics of VHF Signals in the Immediate Vicinity of Trees," *IRE Trans. Antennas and Propagation*, Vol. AP-9, No. 5, pp. 487-491, (September 1961)

LaGrone, A. H., "Height-Gain Measurements at VHF and UHF Behind a Grove of Trees," *IEEE Trans. Broadcasting*, Vol. BC-9, No. 1, pp. 37-54 (February 1963)

Lang, R., A. Schneider, Seker and F. Altman, "UHF Radiowave Propagation Through Forests," CyberCom Technical Report CTR-108-01, CECOM Technical Report 81-0136-4, Contract DAAK80-81-C-0136, (September 1982)

Lang, R., A. Schneider and F. Altman, "Scatter Model for Pulsed Radio Transmission Through Forests," CyberCom Corp., CECOM Contract DAAK80-81-C-0136, IEEE MILCOM '83, Washington, DC, (October 1983)

Lang, R., A. Schneider, F. Truong and S. Altman, "UHF Radiowave Propagation Through Forests," CyberCom Technical Report CTR-108-06, CECOM Technical Report 81-0136-6, Contract DAAK80-81-C-0136, (May 1984)

Murray, J., G. Hufford and J. Adams, "A Guide to Wideband Propagation Measurements and Data Collection-Planning Reference," US Dept. of Commerce, Institute for Telecommunications Sciences, (Revision 5) (December 1984)

Nelson, R. A., "UHF Propagation in Vegetative Media," Final Report, Contract DAAG29-76-D-0100, SRI Project 8998, SRI International, Menlo Park, CA (April 1980)

"Request for Proposal," Wideband Propagation Measurement/Data Acquisition and Recording System, Contract DAAB07-82-C-J097, US Army CECOM, Fort Monmouth, NJ (May 1982)

Rice, P. L., "Some Effects of Buildings and Vegetation on VHF/UHF Propagation," IEEE Mountain West-EMC Conference (1971)

Rino, C. L., J. Owen, G. A. DesBrisay and B. C. Fair, "Calibration and Complex TVIR Construction for the WPMS System," Contract DAAB07-82-C-J097, SRI Project 4520, SRI International, Menlo Park, CA (November 1985)

Rino, C. L., J. Owen, G. A. DesBrisay and B. C. Fair, "The Effects of Antennas in WPMS Channel Diagnostics," Semiannual Report 4, Contract DAAB07-82-C-J097, SRI Project 4520, SRI International, Menlo Park, CA (June 1986)

Rino, C. L., J. Owen and G. A. DesBrisay, "Data Collection Activities Using the WPMS," Equipment, Test, and Demonstration Report, Contract DAAB07-82-C-J097, SRI Project 4520, SRI International, Menlo Park, CA (July 1986)

Sass, P., "Propagation Measurements for UHF Spread Spectrum Mobile Communications," IEEE Transactions on Vehicular Technology, Vol. VT-32, No. 2, (May 1983)

Saxton, J. A. and J. A. Lane, "VHF and UHF Reception-Effects of Trees and Other Obstacles," *Wireless World*, Vol. 61, No.5, pp. 229-232 (May 1955)

Schneider, A., "UHF Radiowave Propagation Through Forested Propagation Channels," NATO AGARD Conference Proceeding 363, on Propagation Influences on Digital Transmission Systems, Athens, Greece, (June 1984)

Schneider, A. and F. Altman, "Proposed Forest Transmission Experiments," CyberCom Technical Report CTR-108-07, Contract DAAK80-81-C-0136, (June 1984)

Staiman, D. and T. Tamir, "Nature and Optimization of the Ground (Lateral) Wave Excited by Submerged Antennas," *Proc. IEE*, Vol. 113, No. 8, pp. 1299-1310 (August 1966)

Tamir, T., "Radio Wave Propagation Along Mixed Pathes in Forest Environments," *IEEE Trans. Antennas and Propagation*, Vol. AP-25, No. 4, pp. 471-477 (July 1977)

Trevor, B., "Ultra-High-Frequency Propagation Through Woods and Underbrush, *RCA Review*, Vol. 5, No. 1, pp. 97-100 (July 1940)

Current Issues in Equalization

Session Chairman: Dennis Hall

John Proakis
Overview of Adaptive Equalization

John Treichler
Adaptive IIR Equalization, Frequency Domain Adaptive Filters

John Cioffi
Algorithms for Multipath, Multitone Equalizers

Brian Agee
Blind Adaptive Equalization

ROBERT SCHOLTZ: Hello, hello .... Have you found me yet? In an effort to get things started on time, I think we'll just let the stragglers miss the beginning. Looks like a little cool weather's blowing in .... If it gets worse, I'll tell you this afternoon where we'll have the cocktail party – we won't cancel it, we'll just move it somewhere a little warmer. I'd like to invite Dennis Hall up now, from TRW, who put together a terrific session for you. Dennis ....

DENNIS HALL: Oh, yeah, my session was very easy to put together – almost everyone volunteered to do it, so it was only a matter of playing phone tag long enough to talk to all of them. John Proakis is here, and he's going to give us a quick overview and then go into some of his current research. We're really fortunate to have John Treichler here; he's got a lot of experience in adaptive filters and equalization. John Cioffi is here, he's exploring some new things in multi-tone equalizers, which I'm very anxious to hear about, and he's done a lot of work in multipath – one of our students that we're sponsoring up to him is looking at that. And I think Brian Agee is going to close it off with reporting some of the continuation of some of his work on blind adaptive signal restoration, which fits well into the modulation characterization session we had earlier.

Let's see, John Proakis – we were lucky that we didn't have to pay him an honorarium to speak with us [LAUGHTER], but you'll notice that his new book is out, or the 2nd edition of his book, so each of us will have to buy 20 copies of his book [LAUGHTER]. John Treichler – John graduated from Rice a couple of years before I did, and in such he'll be the only speaker here who'll speak clearly and without an accent [LAUGHTER]. So that's all I have to open this up, so ....

JOHN PROAKIS: *Overview of Adaptive Equalization*

I should tell you that I did not bring a copy of my book to the Workshop, but one of the people here does have a copy. I found it too heavy to carry, frankly [LAUGHTER]. But Bill Lindsey was not worried about the weight – he did manage to bring a copy.

I'm going to give a brief overview of adaptive equalization techniques. We begin with a general block diagram, shown in FIGURE #1. I've subdivided equalizers into three different categories: linear equalizers, decision feedback equalizers, and equalizers based on maximum likelihood sequence estimation using the Viterbi algorithm. Of the linear equalizers, I'm going to confine my remarks to time domain equalizers, since I know John Treichler will talk about frequency domain equalization. Of the three types that we have listed here, the linear equalizer is by far the simplest to implement and the one that gives the worst performance, basically, in a channel that has significant intersymbol interference. The second type, the decision feedback equalizer, is more powerful as most of you know. The third type, based on maximum likelihood sequence estimation, is the optimum equalization scheme, with one proviso, that we know the probability density function for the underlying statistics of the signal and noise. If we do not know this and attempt to build a maximum likelihood sequence estimation scheme for intersymbol interference, it's quite possible that the performance will not be as good, for example, as the decision feedback equalizer, which does not make any assumptions at all about the statistical properties of the underlying noise process.

Let me first talk about linear equalization in greater detail. Linear equalizers are usually implemented in practice as fractionally

OVERVIEW OF ADAPTIVE EQUALIZATION TECHNIQUES

John G. Proakis
Northeastern University
Department of Electrical and Computer Engineering
360 Huntington Avenue
Boston, Massachusetts 02115

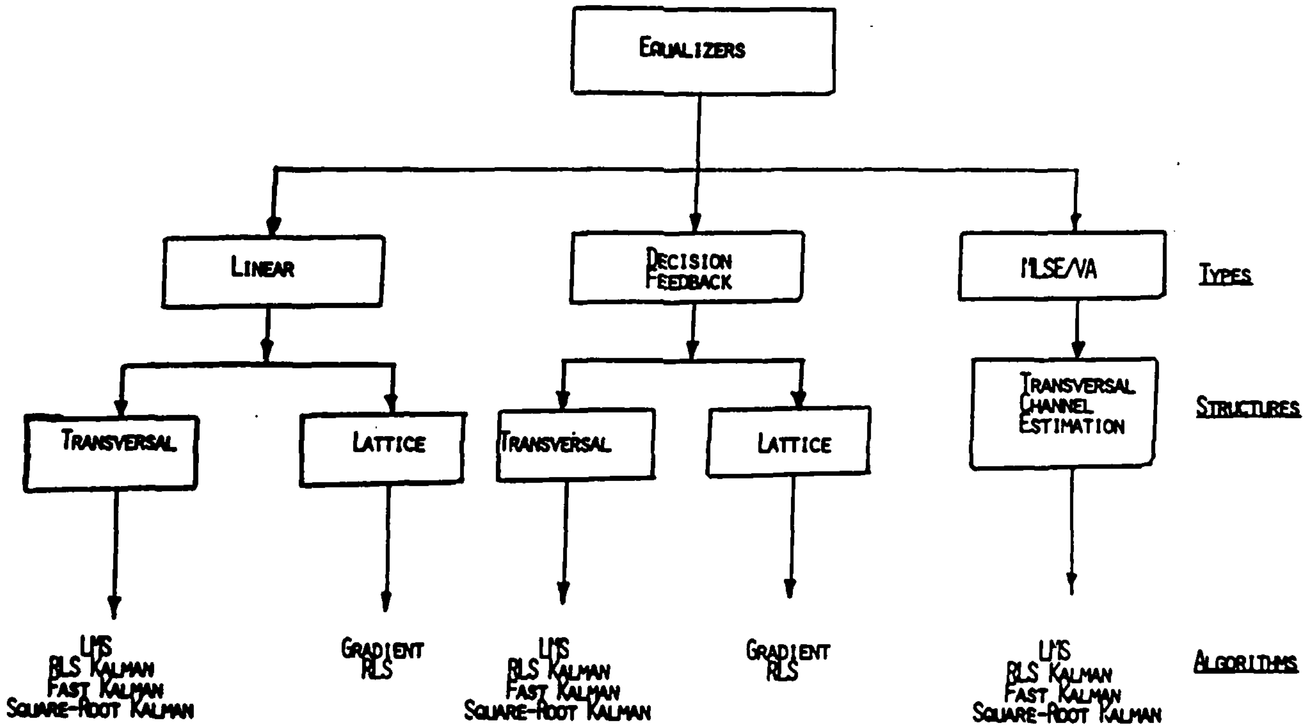


Figure 1 - Types, Structures, and Algorithms for Equalizers

spaced equalizers, and by this I mean that the incoming signal is sampled at least as fast as the Nyquist rate. For example [FIGURE #2], if we have a signal with a raised cosine spectrum and a roll-off factor of β, the highest frequency in the signal is $\frac{1+\beta}{2T}$ where T is the symbol interval. The signal, then, should be sampled at least at twice that frequency, which is $\frac{1+\beta}{T}$, and then passed through an equalizer with tap spacing of $\frac{T}{1+\beta}$. IF $\beta = 1$, for example, the spacing will be $\frac{T}{2}$; if $\beta = 1/2$, we can get by with sparser spacing of $\frac{2T}{3}$ and so forth. In general, then, the spacing is $\frac{M}{TN}$ where both M and N are integers.

The advantage of fractionally spaced equalizers is that in practice we do not know the characteristics of the channel, and the fractionally spaced equalizer plays the role of a matched filter as well as an equalizer. Thus, in one structure we can implement, in an adaptive fashion, a matched filter and a linear equalizer.

A decision feedback equalizer [FIGURE #3] simply uses a linear equalizer as a feedforward filter and then employs the decisions at the output of the decision device in a second filter, called a feedback filter, to cancel out intersymbol interference from previously detected symbols. This particular form is the classical form of the decision feedback equalizer and has some problems if you have coded modulation. In coded modulation you would like to have in your feedback filter the actual decisions from the Viterbi decoder if it's a trellis code, or block decoder if it's a block code. As can be seen, this structure doesn't lend itself very well to coded modulation. For that purpose, there is another structure which is slightly suboptimum called the prediction-type decision-feedback equalizer. I believe that Vedat Eyuboglu is going to say something about that later today, so I will not comment any further on that.

The sequence estimation scheme requires knowledge of the channel characteristics and for this purpose one can use an adaptive channel estimation scheme as indicated in broad terms here [FIGURE #4], and feed the channel characteristics to the Viterbi algorithm to perform the equalization.

Now let's move on to the next level of FIGURE #1 – the different structures for implementing equalizers. In the case of a linear equalizer, there are two basic structures that one can employ: the transversal structure [FIGURE #5] or the lattice structure [FIGURE #6]. There are similar structures for decision feedback equalizers. Associated with the particular structures are a number of different algorithms which are listed in FIGURE #1. I have given the types and the structures, followed by the algorithms that one can apply to these various structures.

In the case of the transversal structure there is the conventional LMS algorithm [FIGURE #5], and then a number of so-called fast algorithms: the RLS Kalman, fast Kalman and square root Kalman. All of these are recursive least squares methods for adjusting the coefficients of the transversal structure. The problem with the transversal structure is that if we want to change the length of the equalizer, we must go back and recompute the coefficients; that is, we must begin anew. To avoid that problem one can go to the lattice structure [FIGURE #6] where we can simply add or delete sections quite easily without affecting the coefficients of the other sections of the filter.

Similar remarks can be made about decision feedback equalizer algorithms, wherein the transversal structure – and by transversal I mean that we have two filters, each of which is transversal, the feedforward and the

2. Fractionally-Spaced Equalizer

Sample the incoming signal at least as fast as the Nyquist rate.

<u>For example</u>: pulse with raised cosine spectrum with a roll-off factor $\beta \Rightarrow F_{max} = (1+\beta)/2T$. The signal can be sampled at rate

$$2F_{max} = \frac{1+\beta}{T}$$

and passed through an equalizer with top spacing

$$\frac{T}{1+\beta}.$$

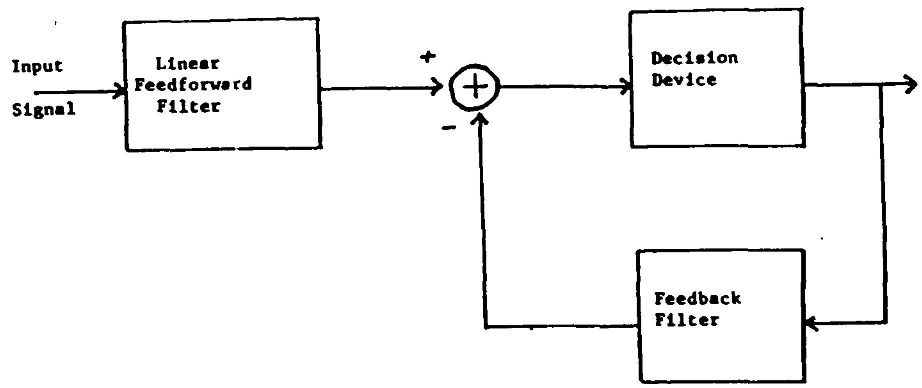
$$\beta = 1 \Rightarrow \frac{T}{2}$$

$$\beta = 1/2 \Rightarrow \frac{2T}{3}$$

etc.

In general, the top spacing may be MT/N where M and N are integers and N > M. Usually, T/2 is used.

Figure 2 - Fractionally-Spaced Equalizers

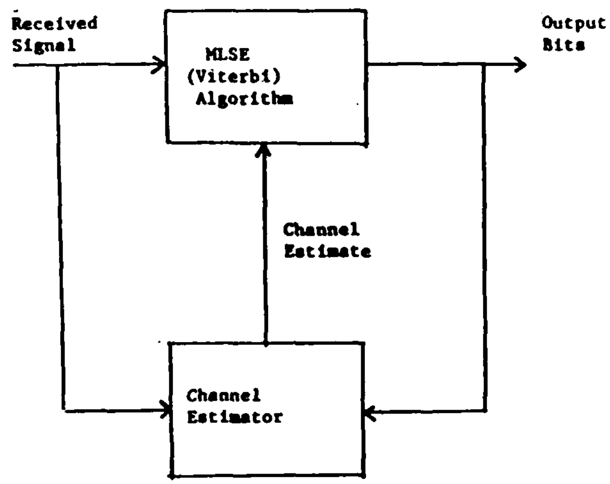


Decision Feedback Equalizer

<u>Note</u>:

There is another type of DFE called the "Predictive DFE"

Figure 3 - Decision-Feedback Equalizer



Adaptive MLSE

Figure 4 - Adaptive MLSE

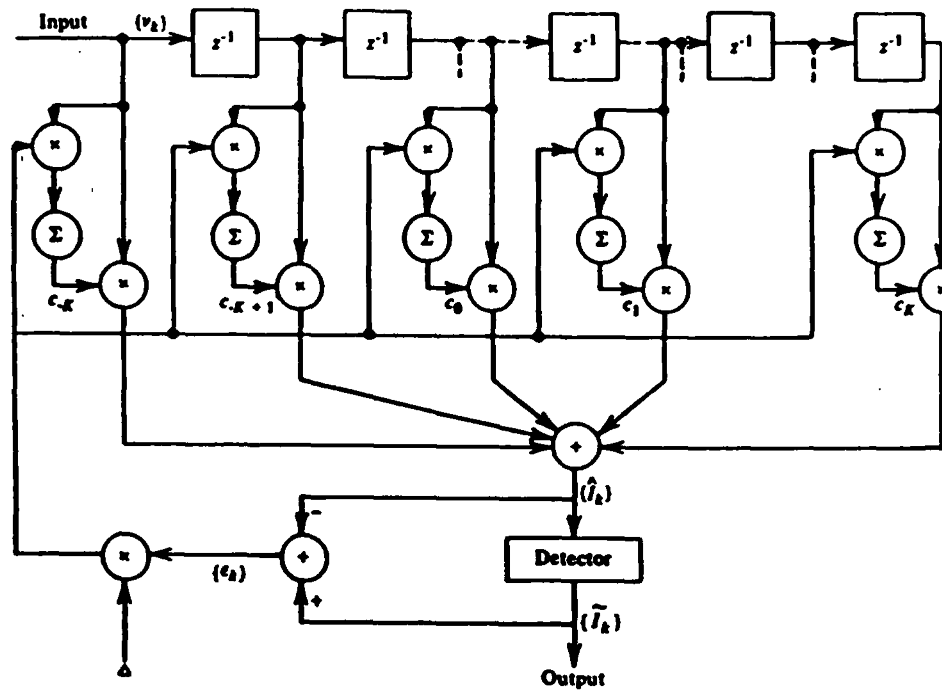


Figure 5 - Linear Equalizer with LMS Algorithm

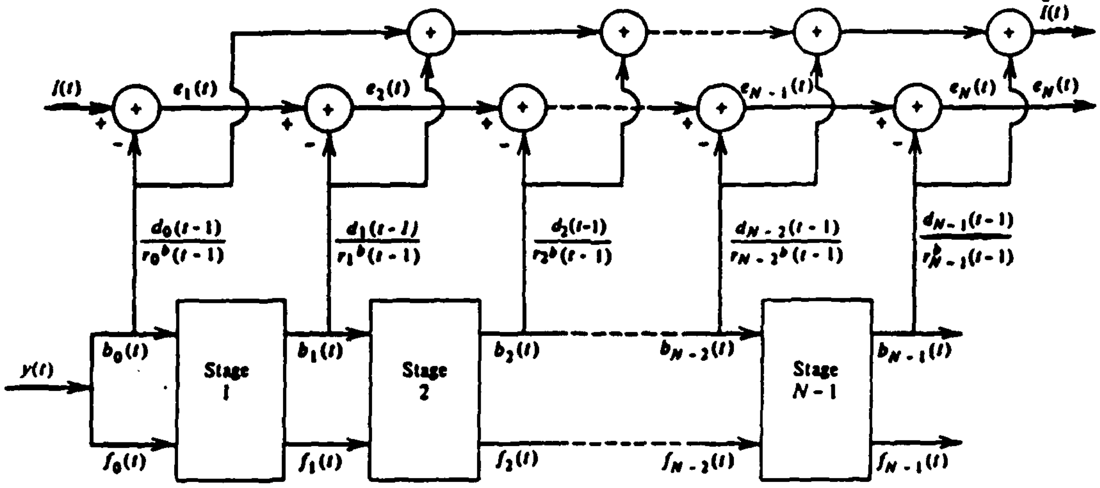


Figure 6 - RLS Lattice Filter

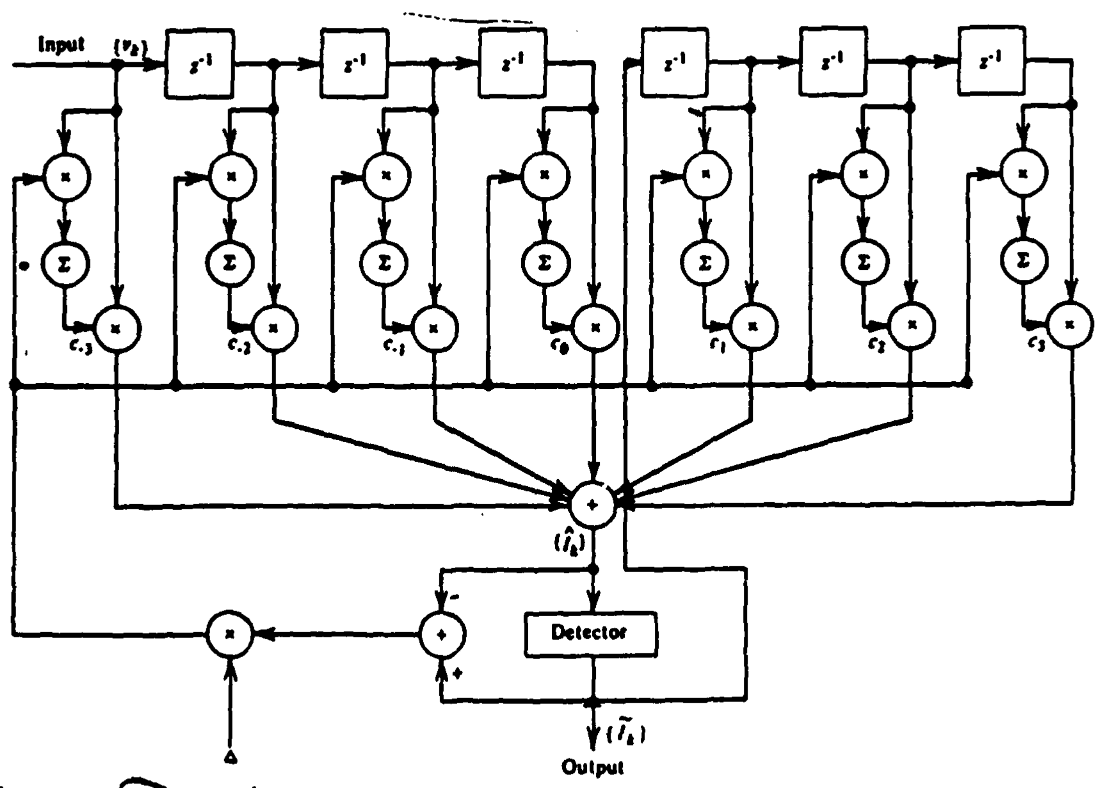


Figure 7 - Decision-feedback Equalizer

feedback filter. One can use the LMS algorithm [FIGURE #7], the Kalman-type fast algorithms, or the square root algorithm for adjusting the coefficients. In the lattice there are two versions of it, the so-called gradient lattice algorithm and a recursive least squares algorithm.

Finally, in the sequence estimation scheme for estimating the channel characteristics, the applicable algorithms are again quite similar since this channel estimator is a transversal-type filter.

Let me briefly describe a recursive least squares lattice for those of you who are not familiar with these lattice-type algorithms. Each of the lattice stages in FIGURE #6 involves a couple of multiplications and a couple of additions and a unit of delay [FIGURE #8]. This is the lattice part. There is also a ladder part which involves multiplying the so-called backward residuals or backward errors by corresponding coefficients and then summing to get the output of the equalizer [FIGURE #6]. This is the input signal, this is the output here, and the desired signal is here. In an adaptive mode, of course, we're feeding back previous decisions or the actual decisions out of the decision device here. A simplification of that is the so-called gradient lattice [FIGURE #9]. The simplification is in a reduction of the computational complexity. The typical lattice stage involves, as I said, multiplying by these reflection coefficients and then adding the signals here. The $b_i(t)$ are the backward errors or the backward residuals and the $f_i(t)$ are the forward residuals in the lattice. In the gradient lattice we force the coefficients in the two arms to be the same, and we update these latter gains via an LMS-type algorithm in an attempt to reduce the computational complexity of the filter. Later, I'll show you some results on the number of computations to implement this type of filter.

A decision feedback lattice equalizer is shown in Figure #10. It seems awfully complicated, but actually it's fairly simple. We have two types of lattice stages. These are the so-called one-dimensional lattice stages and this is a typical stage here, and a number of two-dimensional lattice stages which incorporate both the decisions as well as the incoming signal, and these involve a slightly more complex (two-dimensional type) lattice stage, as shown here. But it is fairly simple in that the input signal comes in here, the output of the decision device is as shown here, and that decision is fed back into this point for computing the various errors in each of the stages of the lattice.

My reason for mentioning these different types of lattice filters is that they do present a number of nice properties over the transversal structures. Let me mention here that in deciding which type of equalizer to use for an application, we need to worry about the convergence rate of the algorithm, the tracking capability, the computational complexity as measured by the number of multiplications and divisions per iteration, the stability of the algorithm, the accuracy, and issues regarding dynamic range if we have to implement these algorithms in fixed-point arithmetic, and scaling problems (see FIGURE #11). Finally, if we wish to implement these in VLSI perhaps we would be interested in considering architectural issues, modularity and parallelism.

The reason for considering the more complex lattice-type algorithms and recursive least squares transversal-type algorithms is simply the fact that they converge faster and, hence, make it possible for us to track time-varying channel statistics. In FIGURE #12,

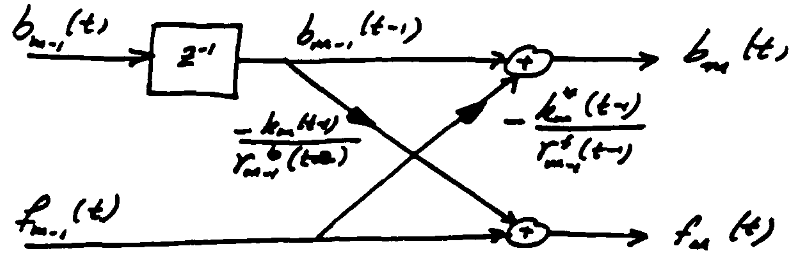


Figure 8- A typical lattice stage

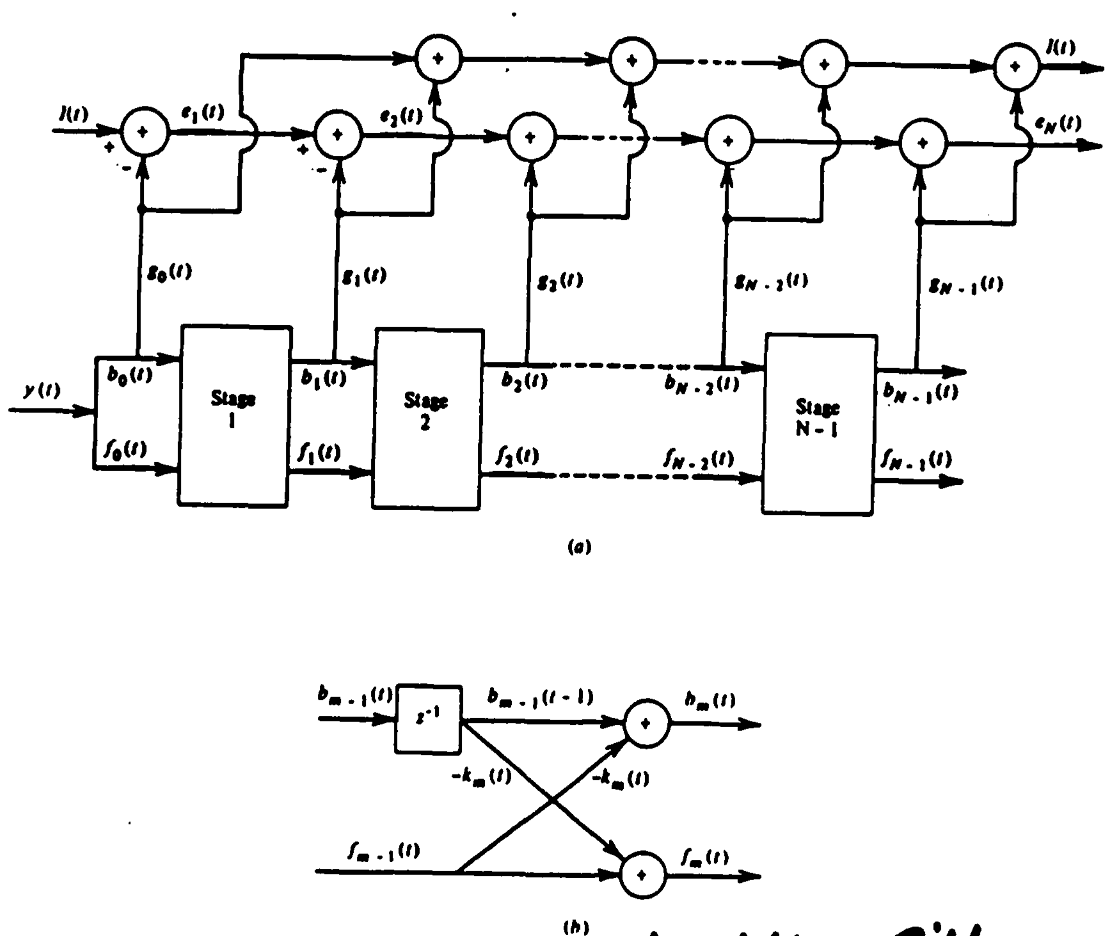


Figure 9- Gradient Lattice Filter

I show the convergence curves for a very bad channel, a channel that has a spectral null. The eigenvalue ratio of the correlation matrix is infinite in this case. These are simulation results, or so-called learning curves, which show the initial convergence of the various equalizers. The fastest converging equalizer is a least squares lattice decision feedback equalizer, which is the one I showed you in FIGURE #10.

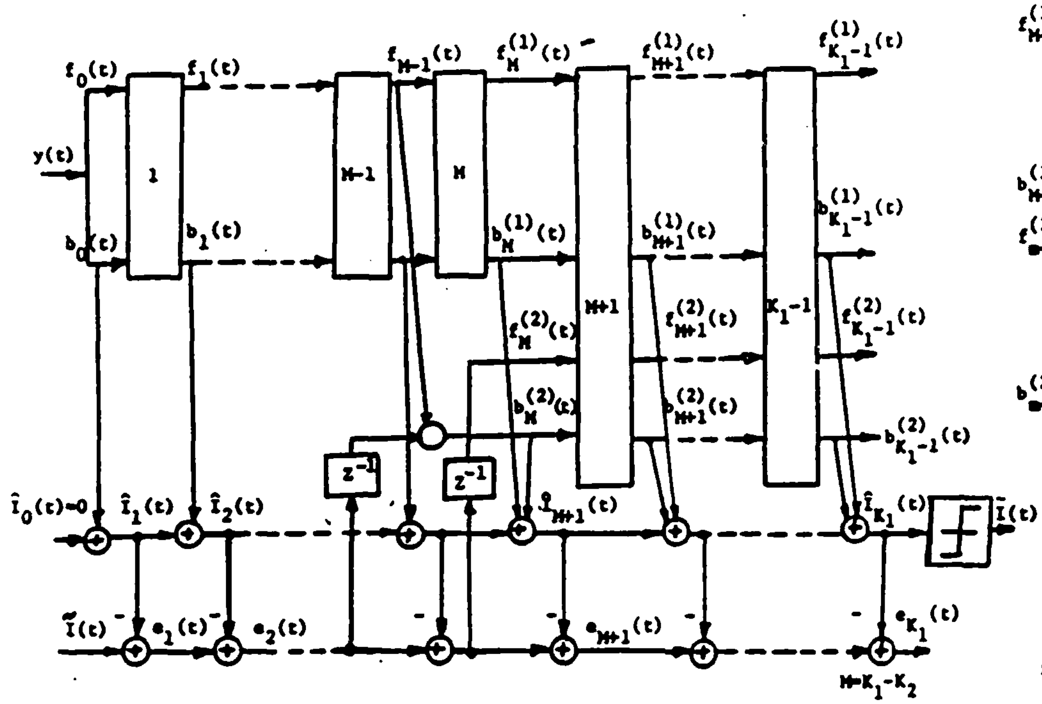
The gradient lattice converges almost as rapidly. It's not an optimum equalizer in the sense of convergence. We give up something in the convergence speed of the equalizer, but we end up with a structure that is less complex computationally. The LMS algorithm, as you can see, converges very, very slowly in this case and takes several hundred more iterations before it reaches this final value. Also I show for comparison the mean square error for a linear equalizer which we know performs very poorly on a channel that has a spectral null. Next is a least squares lattice linear equalizer. It converges relatively fast but then, of course, it bottoms out at a much higher mean square error than the decision feedback equalizer.

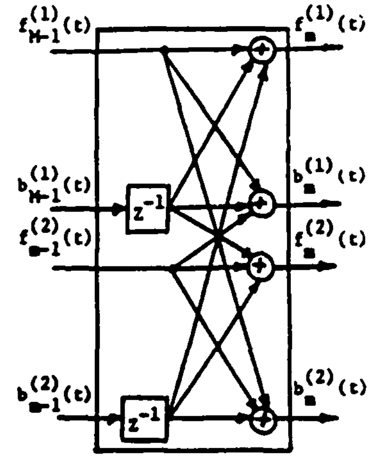
The computational complexity of these different structures is summarized in FIGURE #13, where I plot the number of multiplications (complex multiplications) and divisions as a function of the filter length. As you observe, the LMS algorithm involves the least number of computations. Computational complexity is twice the length of the filter.

The next simplest in terms of computations is the fast RLS algorithm or fast Kalman-type algorithm. I should mention here that this algorithm, at least for the computational complexity which I show here, is based on a form that usually goes unstable after several hundred iterations. This instability is a well known problem with this particular algorithm. Recently, however, a Stanford graduate student by the name of Slock showed how to stabilize this algorithm, but at an increase in the number of computations. So the stable version of this algorithm has a higher slope, but it is still linear in the number of computations. The gradient lattice, as you can see, is the next best in terms of computational complexity, followed by the recursive least squares lattice. And all of these three algorithms are linear in the number of computations as a function of filter length.

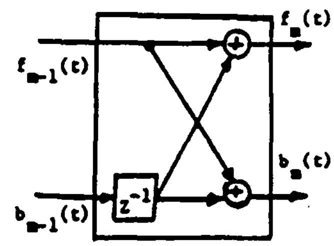
Finally, the so-called square root form of the RLS transversal structure grows as the square of the number of computations. Hence, for short equalizers it's quite competitive with these other schemes, that is, for equalizers of five or fewer taps. But for long equalizers the computational complexity goes up very, very rapidly.

FIGURE #14 shows some results on the numerical accuracy of these algorithms when implemented in fixed-point arithmetic. Sixteen bits is, for all practical purposes, floating point arithmetic, as far as the algorithms are concerned, there's no degradation in performance. I show here the minimum mean square error (which is scaled by 10^{-3}) for the square root algorithm, the fast Kalman or RLS algorithm, two types of lattices, the so-called conventional lattice and an error feedback form of the lattice, and finally the LMS algorithm. We observe that with 13-bit arithmetic there's relatively little degradation in the performance of these various algorithms. With 11 bits the square root algorithm is still performing quite well and the fast RLS is also performing well. The conventional least squares lattice algorithm is quite sensitive to round-off noise and begins to give us poor





2-Dimensional Lattice Stage



1-Dimensional Lattice Stage

Figure 10 - RLS Decision Feedback Lattice Equalizer

Figure 11

PROPERTIES OF ADAPTIVE EQUALIZATION

ALGORITHMS

1. CONVERGENCE RATE AND TRACKING CAPABILITY
2. COMPUTATIONAL COMPLEXITY (Number of multiplications and divisions/iteration)
3. NUMERICAL STABILITY
4. NUMERICAL ACCURACY
5. DYNAMIC RANGE (overflow, scaling)
6. IMPLEMENTATION CONSIDERATIONS (architecture, parallelism, modularity)

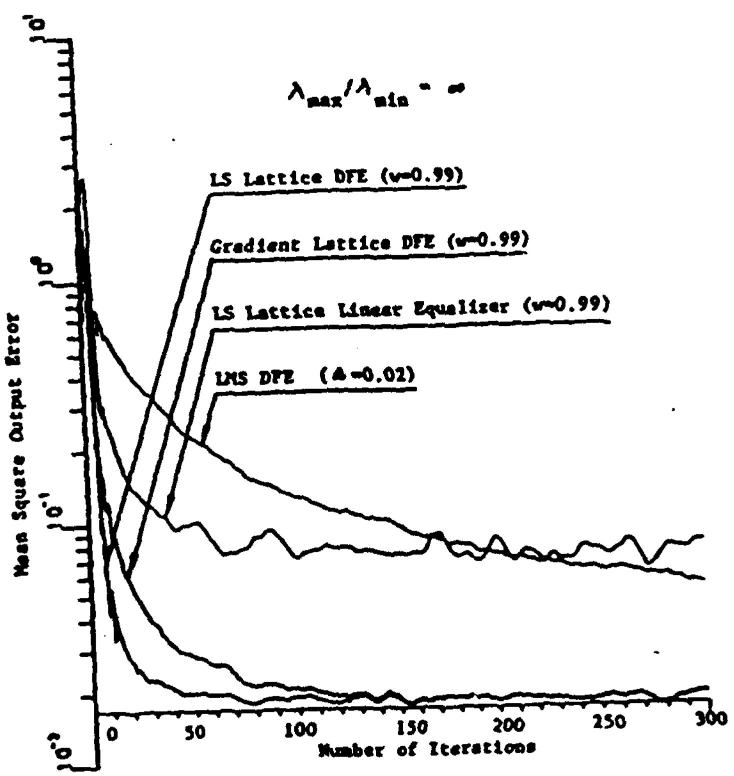


Fig. 12 Convergence Rate of Equalizer (Channel 2)

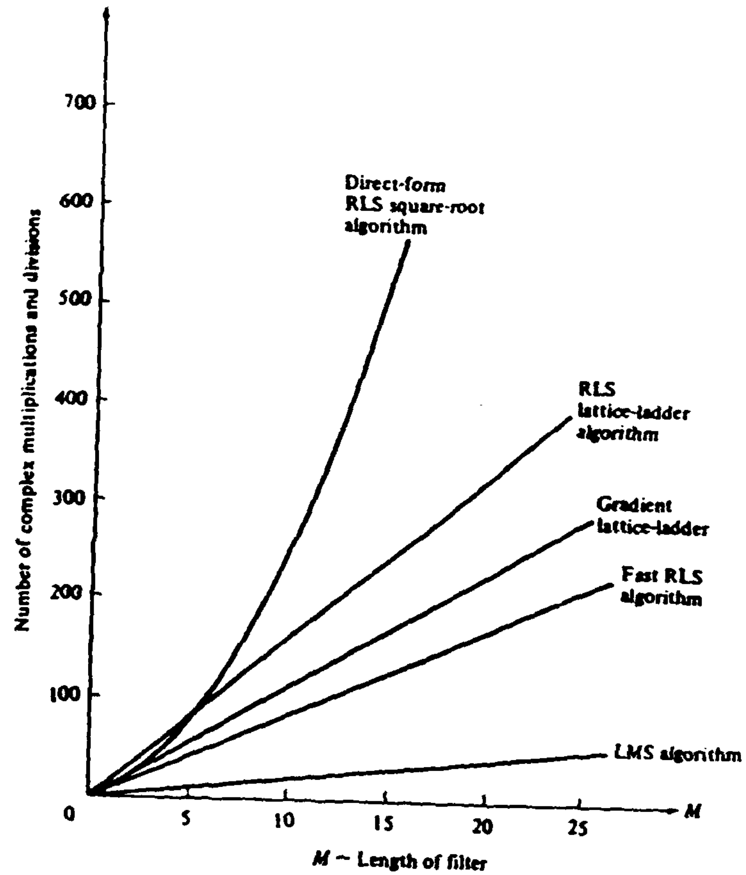


Figure 13 - Computational Complexity

results at this point. The modification to the so-called error feedback form of the lattice gives better performance. The LMS algorithm is also degraded considerably for 11 bits. Finally at 9 bits, we note that the difference between the conventional lattice and the error feedback form is significant. There's a factor of 10 roughly in the minimum mean square error for these algorithms. The fast Kalman algorithm did not even converge in this case, but again if we employ the modification that has been proposed by Slock, I believe that this will take care of the instability.

(PROAKIS IS CONVERSING WITH STEIN, KAWAGUCHI AND OTHERS, WHOSE QUESTIONS ARE NOT DETECTED BY THE MICROPHONE): I'm showing the minimum mean square error scaled by 10^{-3}. Is this your question? [PAUSE] Oh, the eigenvalue ratio here is 11. It's not a bad channel. This is not the same channel that I was dealing with earlier which had a spectral null. These are the number of bits that we carry in all internal computations in the implementation of the equalizer, including the incoming signal which is quantized to 16, 13, 11 and 9 bits, respectively. This was an equalizer of length 10 to 12, so the number of eigenvalues is equal to the number of taps in the equalizer. The ratio of the maximum to the minimum eigenvalue here is 11, which is an indication of how bad or how good the channel is. When we multiply two numbers we truncate to 9 bits, for example. So you can see that all of these algorithms are quite robust and some are more robust than others. In the case of the lattice, the error feedback form is certainly superior and one that gives good performance. Hence, I believe that it's a very nice structure to implement when a fast converging equalizer is required. All of the results that I'm showing you here are in journal papers that we have published, and also in the second editions of my book on *Digital Communications.*

To move to another related subject, I should say something about blind equalization since there is so much interest among the various people here. There are a number of algorithms that have been developed over the last 10 or 15 years on blind equalizers (see FIGURE #15). The first one listed is Sato's Algorithm, which is applicable to PAM, and has been around since 1975. Subsequently we have Godard's paper published in the early 80's, in which he gave an algorithm for QAM-type signals. Later, Benveniste and Goursat proposed some alternative algorithms, which also seem to work quite well. Prati and Picchi give a so-called stop-and-go algorithm for blind equalization and claim that their algorithms work better than the previous three algorithms.

Listed as No. 5 is an algorithm that has come to be known as Bussgang's Algorithm. This was published in a paper by Bellini at a conference two or three years ago. There are some additional algorithms which are still under development, some of which are based on higher order moments of the signal. These are cumulant-type methods which are still under development. Perhaps some people may wish to comment on their experiences with these types of blind equalization algorithms.

Let me conclude by mentioning what my students and I are working on at Northeastern University [FIGURE #16]. I have a student who's working on improving the performance of fractionally-spaced equalizers, basically through the use of singular value decomposition as a means for reducing the excess mean square error, which is generated in fractionally-spaced equalizers. I'm not sure

Figure 14

Numerical Accuracy, in Terms of Output MSE

for Channel with $\lambda_{max}/\lambda_{min}$ = 11 and w = 0.975

MSE x 10^{-3}

Number Bits (including sign)	Algorithm: RLS Square-Root	Fast RLS	Conventional RLS Lattice	Error Feedback RLS Lattice	LMS
16	2.17	2.17	2.16	2.16	2.30
13	2.33	2.21	3.09	2.22	2.30
11	6.14	3.34	25.2	3.09	19.0
9	17.6	*	365.	31.6	311.

*Algorithm Did Not Converge

Figure 15

BLIND EQUALIZATION

1. SATO'S ALGORITHM
2. GODARD'S ALGORITHM
3. BENEVISTE AND GOURSAT'S ALGORITHM
4. PRATI AND PICCHI ALGORITHM
5. BUSSGANG'S ALGORITHM
6. OTHER ALGORITHMS

Figure 16

CURRENT RESEARCH ACTIVITIES

1. IMPROVING THE PERFORMANCE OF FRACTIONALLY SPACED EQUALIZERS
   Excess MSE is proportional to number of nonzero eigenvalues of correlation matrix. Excess MSE can be reduced by use of SVD -> Investigate fast algorithms for SVD.
2. DEVELOPMENT OF NEW FAST ALGORITHMS FOR ADAPTIVE EQUALIZATION AND INVESTIGATION OF THEIR PROPERTIES
3. FAST ALGORITHMS FOR CHANNEL ESTIMATION AND EQUALIZATION OF MOBILE RADIO CHANNELS
4. DEVELOPMENT OF COMPUTATIONALLY EFFICIENT ALGORITHMS FOR BLIND EQUALIZATION
   Use of higher order moment methods
   Improvement of exisiting algorithms

how many of you are familiar with this particular problem, but in a fractionally-spaced equalizer we have a number of zero eigenvalues. In effect, if we implement the equalizer in a conventional way, each tap contributes linearly to the excess mean square error, the misadjustment error that Widrow talks about in the LMS algorithm. What we're trying to do, through singular value decomposition, is to isolate the zero eigenvalues and basically reduce the amount of excess mean square error that is contributed as a result of the fractional tap spacing. The problems are computational in nature because it's computationally difficult to do singular value decomposition in an efficient manner.

Another topic involves the development of new fast algorithms. In spite of what some people may believe, that fast algorithms have already been developed and that there's nothing new to be done here, I will say that this is not the case. It's possible to develop new fast algorithms for adaptive equalization, and a number of these are currently being developed by several people, one of whOM is Fuyun Ling, who is working at Codex Corporation.

We're also working on channel estimation algorithms for equalization of mobile radio channels. I think that's an interesting area for additional work. And, finally, I mention blind equalization as the fourth item here, and the use of higher order moment methods for designing blind equalizers. Also, I believe that there is still some additional work to be done on improving existing algorithms, such as the Godard Algorithm and Benveniste's Algorithm. This concludes my talk.

HALL: Just a little over, OK. Very interesting talk though. Okay, next is John Treichler, who needs no introduction – or has none, either way!

JOHN TREICHLER: *Adaptive IIR Equalization, Frequency Domain Adaptive Filters*

Let me add Item 4c to John's talk, right there [LAUGHTER] ... I can't resist already, and that is that while I think there's definitely room to develop new blind equalization algorithms, I think one of the most crying needs is to understand the ones we've got and how they work, and what performance can be expected of them. They're sort of magic still. The handout you have says that I'm going to talk about frequency domain filters and maybe IIR filters, and stuff like that. Since I've got 20 minutes instead of some much longer interval of time, I thought I'd talk about one of these topics and let the others come later perhaps. I put these all under the category of trying to reduce the amount of computation needed to do adaptive equalization or adaptive interference reduction, and I'll motivate that problem in a moment. It's very nice having John go first because now there's a whole bunch of things I don't need to describe to you – and one of them (all I'm going to talk about) is linear equalizers. I'm not going to talk about decision feedback and I'm not going to talk about mean square error or Viterbi aided methods. I'm going to do straight stuff, and I'm going to worry about how to minimize the amount of computation. Not to say that the others don't have better performance, perhaps, but I always find myself operating up at the very edge of what you can really do in hardware, and you tend to go simple there. Like, how do I minimize the number of computations, what's the simplest possible algorithm I can use as opposed to the best performing?

Generally speaking, the systems that I've been traditionally interested in have some sort of adjustable digital filter [VIEW-

GRAPH #1]. It says FIR, but you don't necessarily need to believe it – it could be something else, could be lattice which is FIR unless it's an IIR lattice. Then what we want to do is take a signal that's been horribly soiled by multipath and/or interference, clean it up, and pass it on to often a traditional signal demodulator or detector. In order to choose the transfer function of that filter, we're going to have some sort of performance measurement and then based on that performance measurement, turn around and adjust the filter coefficients appropriately. I've shown a couple of dotted lines here – maybe they both should be dotted. We either look at this signal (the filter output) or perhaps this signal (the detector output) in order to help us decide whether or not we're happy, and whether or not to adjust the filter coefficients.

The problems I want to talk about today are those characterized by situations where I would like to have an incredibly long impulse response in my filter. Incredible might be, according to John Proakis' pictures, 13. But I was really thinking like hundreds or thousands of taps. Situations where I'm trying to receive, for example, a high bandwidth signal that requires a high sampling rate in the presence of fairly considerable multipath, impulse responses that last say microseconds. Microseconds aren't very much if my sampling rates are kilohertz, but if my sampling rate has to be, say, a hundred megahertz, I've got a hundred tap filter, or maybe a 200 or 300 tap filter. Other situations where a long impulse response is really important is if I'm trying to excise narrowband interference. Think back yesterday to Phil Bello's talk about where I've got a measly megahertz of bandwidth, but I have all these interferers in it, and I would like to chop them out. In fact the resolution I would like to have is maybe a thousand times or maybe even 10,000 times the sampling rate. Is it possible to ever get there?

Now there are several different methods, and as I mentioned, given the fact that I promised up and down that this is only going to take 20 minutes, I'm only going to talk about one of these three areas [VIEWGRAPH #2]. I would be happy to have somebody ask me questions later in the day or tomorrow so I can talk about the others. There have been three basic approaches that have come along for handling this. All have different advantages and different amounts of success, and different amounts of attention that have been paid to them. All of them are linear equalizers. The first one which I am going to talk about is frequency domain filtering where we try to use the FFT to help us. Others are adaptive IIR filtering, and that sounds like a good idea, too. Hey, if I use IIR filters I can get long impulse responses without much computation, and that's true ... but you also get some other stuff that you may not have wanted. Okay, then the last technique is to use sparse filters where I have a very long tap delay line, but don't bother to put all the coefficients in. That turns out to have some nice applications where the filter needs to be really long and densely tapped because of the high sampling rate. I'd like to poke out a lot of interferers, I mean use a lot of notches, in the transfer function. But a lot might be only a hundred, where in fact I may need a thousand taps in order to get the resolution I want. So are there games that I can play where the degrees of freedom I need are much fewer that the number of taps I need to gain the resolution I want? Now that I've teased you about it, I'm not going to talk about it!

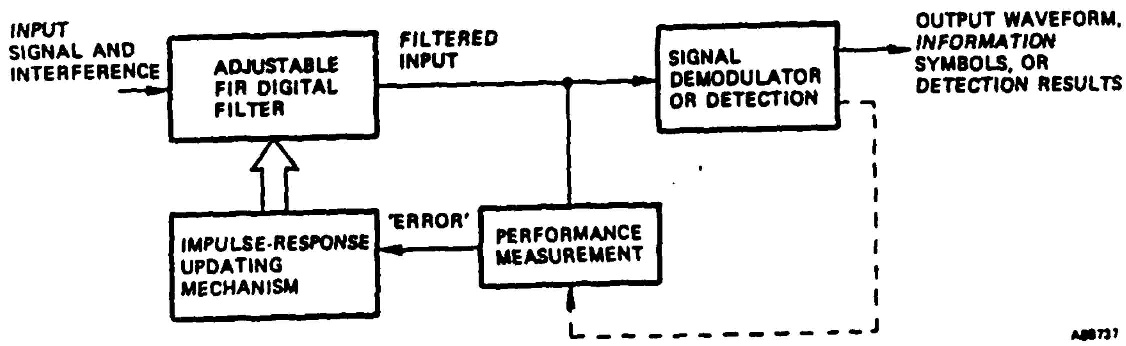
Okay, what I want to talk about today, at least in this interval, is frequency domain

REDUCING THE COMPUTATION
NEEDED FOR
ADAPTIVE INTERFERENCE SUPPRESSION

B-282-89

SYSTEM OBJECTIVE

- DEVELOP A LONG-DURATION, LOW-COMPUTATION ADAPTIVE FILTER CAPABLE OF MITIGATING TIME-VARYING MULTIPATH OR OTHER ADDITIVE INTERFERENCE



VIEWGRAPH #1

B-282-89

adaptive filtering [VIEWGRAPH #3]. I'm going to try to restrain myself and not tell you what limited I know about it, but just some of the stuff at the top: what people are doing, what appears to work, where the problems are, what you can hope from it, and so forth. I've started all these three sections, only one of which I'm going to show you, with the promise – you know, this is magic. Here's what I'm going to go to NSF with, or Office of Naval Research or ARO, and say, "Here's the big deal, and I want a whole lot of money to study this." Okay, the promise is that I ought to be able to reduce the amount of computation quite considerably for high-order high-resolution FIR filters. Why? Because the FFT gives me, you know, log N as opposed to N, and it's just natural. Now the other thing that's more subtle, but turns out to be just as good an argument for the frequency domain techniques, is that it gives me an opportunity to orthogonalize the input. Well, not for all signals, but it turns out lots and lots of the signals that we practically care about might be characterized as relatively broadband, with a whole lot of narrowband interference. You know about Karhuren-Loève and all that, and I can orthogonalize it a bunch of different ways, but it turns out that if my input is well characterized by a whole bunch of narrowband signals, then an FFT is a very nice way to somehow orthogonalize or separate or channelize all those different pieces. And that may allow to do something neat, and we'll talk about that in a minute – it has to do with eigenvalues and such.

OK, successes, does anybody really use this stuff? Yes, these methods have already been proven, unfortunately in a few applications I can't talk about broadly. But the people have built these things and they work very, very well in applications where you have relatively weak broadband signals of interest in the presence of a tremendous number of time-varying narrowband interferers, and in the presence of multipath and in the presence of other things where I would like to be able to use this long equalizer. There are basically two technical approaches; I'm going to focus on one of them today, and I'm going to explain to you why I focus on that one [VIEWGRAPH #4]. The first one is the straight use of fast convolution and fast correlation. If you think back to the block diagram, there basically has to be in every adaptive filter, a filter, right up here, some method of deciding whether the performance is good or bad, and then some method of taking that error and figuring out what set of impulse responses I would like to use in the filter. If I'm interested in using a gradient descent technique, like LMS or Godard or CMA or any of these other things that we've talked about lately, then I can do the filtering with just pure Oppenheim and Schafer fast convolution – I guess I should actually go back to Tom Stockham and those guys, but the book you ought to see is the orange one – whereas zero-fill and FFT – and this is just a filter – and I tell it the impulse response, and low and behold with log N improvement, I've built a filter. I can build any FIR filter consistent with the length of the FFTs. Similarly if I send in an error signal, and a delayed input, I can use exactly the same techniques through the conjugation to correlate the error with the input, and end up with an estimate of the gradient. Using only "fast convolution" and "fast correlation" I have, with no mess whatever, a rather straightforward implementation of an adaptive filter gradient descent-type in the frequency domain. And that works, it works very nicely.

OK, there is another method that's come

AST

ALTERNATIVE APPROACHES

- FREQUENCY DOMAIN FILTERING
- ADAPTIVE IIR FILTERING
- SPARSE FIR FILTERING

VIEWGRAPH #2

B-282-89

AST

FREQUENCY DOMAIN ADAPTIVE FILTERING

- THE PROMISE
  - LOWER COMPUTATION REQUIRED FOR HIGH-ORDER, HIGH RESOLUTION FIR FILTERS
  - IMPROVED CONVERGENCE SPEED ATTAINABLE BY ORTHOGONALIZING THE INPUT SIGNAL
- SUCCESSES
  - VERY SUCCESSFUL IN RECOVERING WEAK BROADBAND SIGNALS IN THE PRESENCE OF TIME-VARYING, NARROWBAND INTERFERENCE
- TWO TECHNICAL APPROACHES
  - FAST CONVOLUTION/FAST CORRELATION USING THE FFT
  - "FILTER BANK" OR TRANSMULTIPLEXER-BASED CHANNELIZED PROCESSING, ALSO USING THE FFT

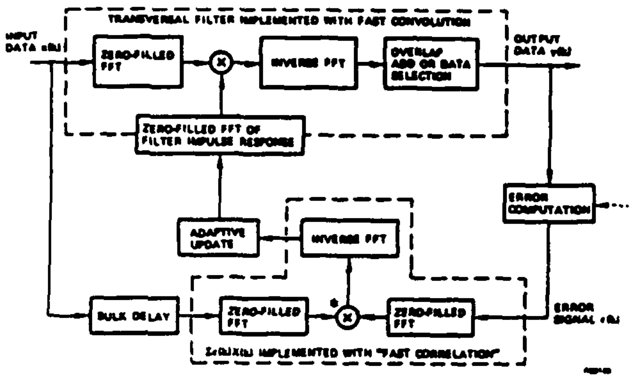
VIEWGRAPH #3

B-282-89

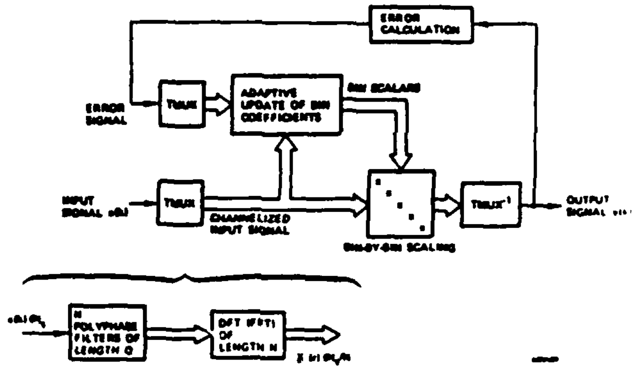
AST

TWO DESIGN ALTERNATIVES FOR FREQUENCY-DOMAIN FILTERING

FAST CONVOLUTION/CORRELATION METHOD



TRANSMULTIPEXED–BASED METHOD



VIEWGRAPH #4

B-282-89

AST

RELATIVE ATTRIBUTES

- BOTH METHODS REQUIRE SUBSTANTIALLY LESS COMPUTATION THAN TIME-DOMAIN FIR SINGLE-RATE FILTERING
- TRANSMUX METHOD USES MORE FILTERING WHILE "FAST" METHOD USES MORE FFTs
- TRANSPORT DELAY IN TRANSMUX-BASED FILTERS IS MUCH LONGER THAN "FAST" DESIGNS AND "FAST" DESIGNS INTRODUCE MORE DELAY THAN THE EQUIVALENT TIME-DOMAIN FILTER
- IF SIGNAL QUALITY ASSESSMENT (I.E., ERROR GENERATION) IS A LINEAR OPERATION, THEN ADAPTIVE UPDATING CAN BE DONE IN THE FREQUENCY DOMAIN FOR BOTH APPROACHES, YIELDING VERY EFFICIENT DESIGNS
- IF ERROR GENERATION IS NOT LINEAR (E.G., DECISION-DIRECTION) THE PROBLEMS BEGIN

VIEWGRAPH #5

B-282-89

along lately that has even, apparently, more advantages. It uses what some of us have traditionally called "transmultiplexers". This is a name used in the multichannel telephone world, mostly, to describe a technique from going from FDM to TDM, or from the frequency domain channelizing of the input into a whole bunch of equal sized frequency channels and then reporting the samples out in sequence or in a time division form. I don't want to go into this too much here, but basically they're built with a set of polyphase filters – the name I hate, but that's what everybody calls them – followed by a DFT, which is usually implemented as an FFT. The combination of these gets you the same log N improvement as before, but has the very nice property that instead of just doing an FFT that has fairly sloppy $\frac{\sin Nx}{\sin x}$ frequency response, depending on how long I make these subfilters here, these polyphase filters, I can make brick wall, beautiful channelized filters. Instead of being down only 13 dB in the next bin, I can be down 60, 80, it all depends on what I want Q to be and how much arithmetic accuracy I'm willing to provide up there. It also has the subtle advantage that, now I'm working in blocks. This data comes in, say I build a transmultiplexer that has a thousand point FFT, and break the signal into a thousand bins, well it turns out that the sampling rate here (at the channelizer output) is now 1/1000 of the input rate. So whatever I deal with here on a channel-by-channel basis is at a decimated rate, and I can do some very nice things.

STEWART LINSKY: Why can't you get the same thing by windowing the input on the FFT?

TREICHLER: You can.

LINSKY: OK. Is that the only advantages of this other program?

TREICHLER: Let me follow that up immediately – these two methods aren't completely distinct. I can start playing all sorts of games back and forth between the two. I've made them look as separate as possible for discussion purposes. I don't want to tell you by any means that this is it, and these are the only ways to do it. But I do want to show you what some of the significant differences are, and I've chosen the poles of no weighting and extraordinary weighting, in order to give you an idea of what can help and what can hurt.

OK, the other thing I've done is with my adaptive filter. I dechannelize, I multiply each channel output – and what I've shown here is a diagonal matrix – is a bin-by-bin or a channel-by-channel weighting with no cross channel weighting going into what I call an inverse transmultiplexer, which you may *think of as just an interpolator*, and an upconverter. All it does is put all these bins back together into a time domain waveform at the same sampling rate as we went in. So I've channelized the signal, I've scaled it on a channel-by-channel basis, and I've put it back together. If this were an analog system these things would be at the same "sampling rate", if you will, as the input, and I wouldn't have had to interpolate; but since I've decimated here in the digital world, I have to interpolate. Now, notice something I've done here – and you may ask why, and I'll elaborate on this – I've shown the output coming from the time domain waveform all the way around through an error calculation and back in. Now, why do I need to do that? Well, I'll have you hold that thought for a moment. Let me compare these two methods for just a minute. Again, I've chosen to make them as different as possible. But what do they have in common, relative attributes? [VIEWGRAPH #5] Both of them have sub-

stantially less computation, and to tell you exactly how much less computation than the straight time domain, you have to tell me what you really want to do. I've shown one particular method of doing the fast convolution/fast correlation method; in fact there are dozens. There's a very nice paper that Clark, Mitra and Sid Parker did four or five years ago on whole bunches of different ways of doing that using 3, 4 or 5 FFTs and different sorts of assumptions. Similarly, in the transmux method, or the channelized method, you have to tell me how big Q is and how big the FFT is. Nonetheless, no matter how you do it, it's less. Like typically an order of magnitude or two less, so that's good. People argue which is more and which is less computation between the two, and again I can change the rules, I can change the assumptions, and I can make them relatively better or worse than each other. Something that we should note, however, is that the transport delay through the thing happens to be much, much more with the transmultiplexer-based filter that has the very nice stopband performance. Similarly if I want to start windowing or weighting the fast convolution method, I end up having to bring more data in and introducing more transport delay there. The implications of that will become clear later.

One of the most important things is my signal quality assessment. For example, those of you who take EE 373 at Stanford University from Bernie Widrow, all he ever talks about the way I form an error, is I subtract the output of the filter from a God-given desired signal, and I've got an error, and I go and adapt with it. That's a very nice, straightforward linear operation, and if God always gave us that desired signal to work with, we'd be in really good shape. If in fact it were there, then nice things happen because it's linear, that operation commutes with the frequency transformation process, and I can do all the updating in the frequency domain. You say that sounds nice; well, it turns out that it's not just nice, it's really important. For example, what if our error computation or our performance assessment method isn't linear? Say, ah well, you're talking about a new algorithm, like those based on cumulants, you're talking about CMA, Godard – I'm talking the more conventional decision direction method. The only way we currently know how to do decision direction is to get filter-output waveform and look at it IQ, and see how close it is at the baud intervals to the constellation points we would like to match, take that difference and go on. It's very much a time domain operation. That's why that feedback loop I showed you from the output all the way back around exists.

Now let's look at a couple of observations. One of them is this nonlinear performance measurement, like decision direction, Godard or CMA. It doesn't commute, you have to use the loop or some better method, and there would appear to be no problem. The transport delay, I pointed out, is bigger. The most important thing is people actually built this thing expecting wonderful performance. Does it channelize wonderfully? Yes. Does it adapt really fast, like they expected it to? No! And the question is why. Well, if you go back to John Proakis' viewgraph, where he showed about eigenvalue ratios, the trick with LMS is that you look at the covariance matrix, and you look at the biggest eigenvalue and the smallest eigenvalue, and the ratio of those tells you a lower bound on the number of iterations it will take for that filter to converge, depending on how you define "converge". And then if you use an adaptation coefficient that isn't full-out, you know, all

the way against the firewall, it's even slower. So this problem with eigenvalue disparity has been a killer. Now if the input is mostly dominated by interference, and if that interference can be channelized, and each one of those things falls in a separate bin, then your eyes start opening wide, and you say, "I can use separate μ's in every bin, and I can adjust those μ's to optimize performance in each bin separately, and I can make each one converge in only one step." Well, you can't! [VIEWGRAPH #6] Here's why. If there were no delay in that feedback loop, then what happens is that the adaptation behavior in each bin can be characterized as a single pole that lies somewhere between here [$z = 1$], where μ is zero, here [$z = 0$] where μ is such that μ times the eigenvalue λ is 1 – and if I keep cranking up μ I can drive it unstable. Now it turns out that as soon as I start using these transmultiplexers in order to get these very nice adjacent channel rejections [VIEWGRAPH #7], I start introducing delay in the loop and I start putting additional poles at the center. As I start cranking up μ – and remember I want to put it here [$z = 0$] – what happens is these guys do the old root locus-spider spread, OK – that's Texas [LAUGHTER] – anyway, it's actually a Texas football play, everybody runs in all directions hoping to catch a long pass. Anyway, what happens is this pole comes in from $z = 1$, and instead of getting anywhere close to $z = 0$, a pole comes out, these scatter, these join and spread, and you go unstable or oscillatory for much, much lower values of μ than you would have ever suspected. [OK, two minutes, right? One minute? One and a half! OK, so the bottom line – I ain't quite to the bottom line, but I'm getting close [LAUGHTER] – this is really good, no sweat!] [VIEWGRAPH #8] Say I had set μ so that $\mu\lambda$ was a tenth – that means that it ought to adapt in ten iterations, which is a nice, safe number – and here's what happens [$m = 1$ curve]. But if I keep μ the same value and start making my transmultiplexers better and better, attaining better and better adjacent channel rejection, here's what starts happening to the transient response of that single coefficient. Instead of converging in say 20 iterations, it can take 80, 90, 100 and worse. (See curves for $m = 3$, 5, and 7.) The better I make the adjacent channel rejection in my bins, the worse the transient performance is. So, you're dead meat, right? Well, yeah. My final viewgraph [VIEWGRAPH #9] – what this means is you've got to start looking at alternatives. That alone isn't going to get it for you. And what do you do? Maybe I ought to abandon the gradient descent and try to use some closed-form methods or accelerated techniques so I get my improved performance some other way. Heck, maybe I look at this as a control problem and try to use feedforward compensation; you know that's a fairly established method, that might work. Maybe I want to – the last two methods, [sorry Dennis] – the last two methods, I do have a viewgraph for, and they are to play games, start playing games. [VIEWGRAPH #10] Here is my original filter. Maybe I want to put something inside the loop that just looks at the spectrum, and tries to respond instantaneously and very rapidly to the spectrum. Like if I see an interferer pop up, hammer it down, and then come around on a slower, known to be slower, basis and clean up the mess. Another method is to somehow try to approximate all this stuff, and if I can get the approximation inside of the two transmultiplexers, the fact that it's an approximation may not hurt me as bad as the improvement I get by being able to do everything on a bin-by-bin basis.

AST

PROBLEMS IN COMPUTATIONAL PARADISE

- **OBSERVATIONS**
  - NON-LINEAR PERFORMANCE MEASUREMENTS (E.G., CMA AND/OR DECISION-DIRECTION) DO NOT COMMUTE WITH FREQUENCY TRANSFORMATION OPERATIONS
    - → ERROR SENSING AND WEIGHT ADAPTATION CANNOT BE DONE PURELY IN THE FREQUENCY DOMAIN; CONVERSIONS INTO AND OUT OF THE TIME DOMAIN ARE REQUIRED TO SENSE THE ERROR
  - THE TRANSPORT DELAY OF A TRANSMUX-BASED FILTER IS APPROXIMATELY QN SAMPLES vs. ABOUT 2N FOR A FAST CONVOLUTION FILTER vs. ABOUT N/2 FOR AN N-POINT TIME-DOMAIN FILTER
  - OBSERVED CONVERGENCE RATE PERFORMANCE OF THE CLOSED-LOOP TRANSMUX-BASED FILTER IS FAR SLOWER THAN ANTICIPATED
- **QUESTIONS**
  - IS THE SLOW CONVERGENCE CONGENITAL?
  - IS IT RELATED TO THE OTHER OBJECTIONS?
  - IS THERE A CURE?

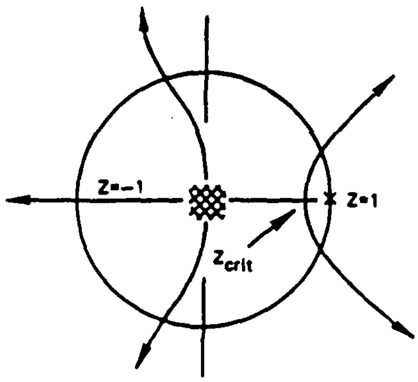
VIEWGRAPH #6

B-282-89

AST

ROOT LOCUS FOR POLES OF ADAPTATION PROCESS AS FUNCTION OF $\mu\lambda$

- THE EXPECTED TRANSIENT RESPONSE OF THE ADAPTIVE CONVERGENCE PROCESS $w_n(z)$ CAN BE ANALYZED USING ROOT LOCUS TECHNIQUES



POLE BEHAVIOR AS A FUNCTION OF $\mu_n\lambda_n$ FOR $\Delta_{tm} = 2$. IF $\Delta_{tm} = 2$, $m \equiv 2\Delta_{tm} + 1 = 5$

EQUATIONS:

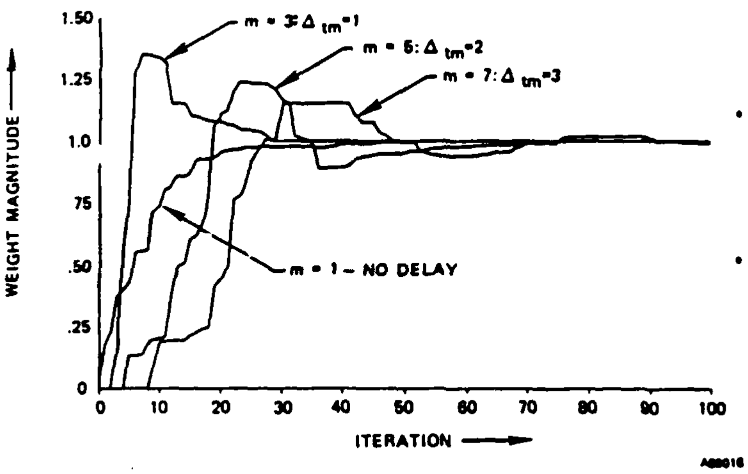
- $z_{crit} = \dfrac{m-1}{m} = \dfrac{2\Delta_{tm}}{2\Delta_{tm}+1}$
- $(\mu_n \lambda_n)_{crit} = \dfrac{(m-1)}{m^m}$
- STABILITY LIMIT $\approx 4.27 \cdot (\mu_n \lambda_n)_{crit}$

VIEWGRAPH #7

B-282-89

AST — TRANSIENT PERFORMANCE AS A FUNCTION OF μ, λ, AND LOOP DELAY

MOVEMENT OF THE POLES CAUSES OSCILLATORY BEHAVIOR AND ULTIMATELY INSTABILITY



WEIGHT BEHAVIOR AS A FUNCTION OF LOOP DELAY FOR $\mu_n \lambda_n = 0.1$

OBSERVATIONS

- WITH NO DELAY, μ CAN BE SELECTED TO MAKE $\mu\lambda$ ALMOST UNITY AND ATTAIN VERY FAST CONVERGENCE
- EVEN SHORT DELAYS FORCE μ TO BE REDUCED SUBSTANTIALLY, THUS GIVING UP MUCH OF THE TOUTED CONVERGENCE RATE ADVANTAGE
- LARGER Q PROVIDES BETTER BIN-TO-BIN ISOLATION BUT FORCES μ TO BE LOWER YET

VIEWGRAPH #8

B-282-89

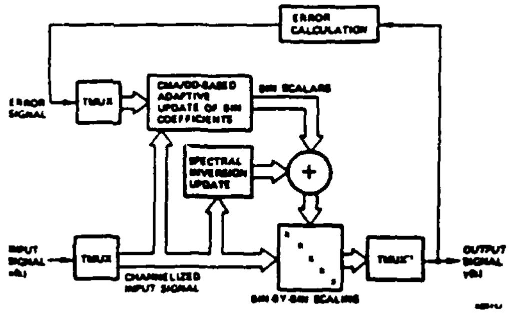
AST — APPROACHES TO SPEEDING UP CONVERGENCE

- ABANDON GRADIENT-DESCENT AND USE CLOSED-FORM OR ACCELERATED GRADIENT TECHNIQUES
- USE FEED FORWARD COMPENSATION IN ADAPTATION LOOP TO IMPROVE TRANSIENT PERFORMANCE
- USE SPECTRAL INVERSION TECHNIQUES TO "PRE-EMPHASIZE" FREQUENCY RESPONSE UNTIL CLOSED-LOOP OPERATION CAN BE EFFECTIVE
- FIND COMPUTATIONALLY SIMPLER FREQUENCY-DOMAIN APPROXIMATION TO ERROR GENERATION PROCESS

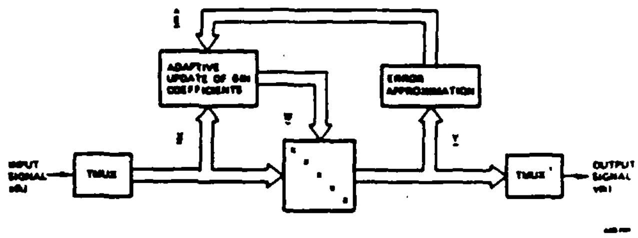
VIEWGRAPH #9

AST TWO OF THE POSSIBLE APPROACHES

INNER LOOP
SPECTRAL INVERSION



FREQUENCY - DOMAIN
ERROR APPROXIMATION



VIEWGRAPH #10

B-282-89

AST IIR ADAPTIVE FILTERING

- PROMISE
  - REDUCE COMPUTATION CONSIDERABLY WHEN MODELING RESONANCES OR COMPENSATING FOR SPECTRAL NULLS
  - REDUCED DIMENSIONALITY OF FILTER MIGHT IMPROVE CONVERGENCE RATES OF ADAPTIVE PROCESS
- ANALYTICAL WORK
  - STRAIGHTFORWARD EXTENSION OF LMS – FEINTUCH, HORVATH
  - RECURSIVE GRADIENT ESTIMATION – STEARNS, WHITE, AHMED
  - ARMA AND ARMAX METHODS – MORF, FRIEDLANDER
  - STABILITY-BASED METHODS – JOHNSON, ET. AL.

VIEWGRAPH #11

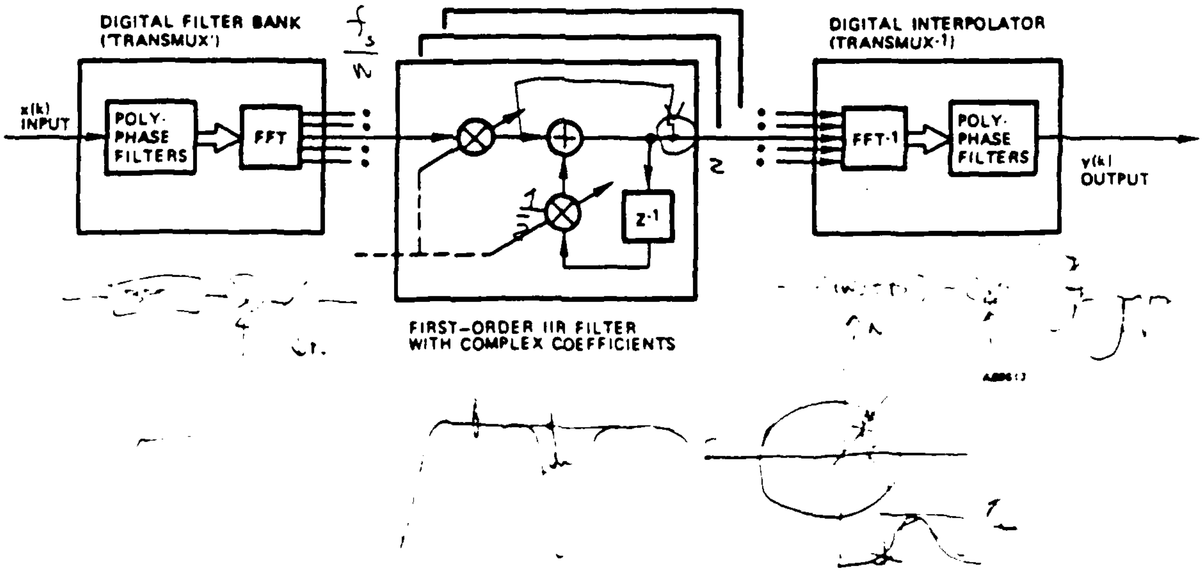
AST IIR ADAPTIVE FILTERING (CONTINUED)

- PRACTICAL SUCCESSES
  - CONSTRAINED DESIGNS
    - POLES LOCKED TO ZERO (THOMPSON)
    - ONLY TWO POLES (CCITT ADPCM ALGORITHM)
  - 1-POLE DESIGNS IN FREQUENCY DOMAIN ADAPTIVE FILTERS
- ANALYTICAL AND PRACTICAL PROBLEMS
  - HIGHER ORDER FILTERS HAVE SLOW AND/OR UNCERTAIN CONVERGENCE
  - CONSTRAINTS/TESTS TO ENSURE STABILITY ARE DIFFICULT AND HARD TO IMPLEMENT
  - EQUALIZATION OF NON-MINIMUM-PHASE CHANNELS REQUIRES UNSTABLE POLE/ZERO CHOICES

VIEWGRAPH #12

B-282-89

AST USING 1-POLE IIR FILTERS IN TRANSMUX-BASED INTERFERENCE "CANCELERS"



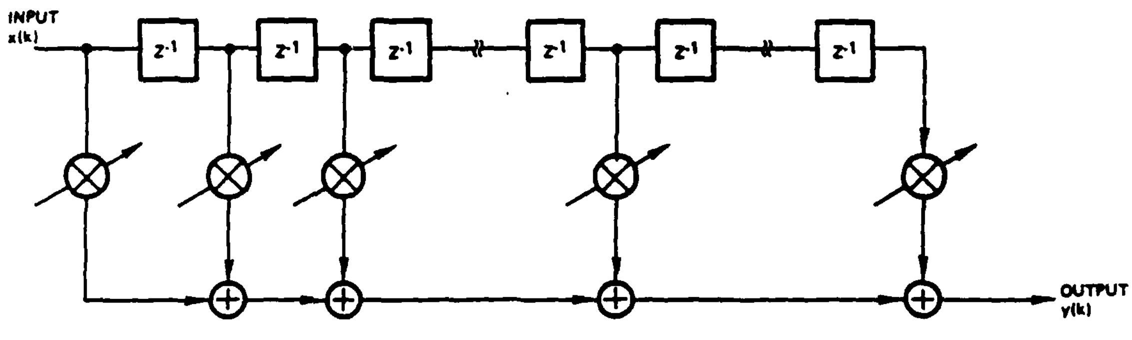
VIEWGRAPH #13

B-282-89

AST

SPARSE FIR TIME-DOMAIN FILTERING

- APPROACH – THIN OUT THE TAPS OF A LONG FIR ADAPTIVE FILTER



- POTENTIAL APPLICATIONS – THOSE WHERE DEGREES OF FREEDOM ARE MUCH LESS THAN THE LENGTH REQUIRED TO ATTAIN THE DESIRED RESOLUTION

VIEWGRAPH #14

B-282-89

AST

SPARSE FIR TIME-DOMAIN FILTERING (CONTINUED)

- FAVORABLE ATTRIBUTES
  - SIMPLE EXTENSION OF STANDARD FIR TIME-DOMAIN DESIGN
  - GRADIENT-SEARCH ALGORITHMS (E.G., CMA AND LMS) WORK WITH NO CHANGE
- TECHNICAL RESEARCH ISSUES
  - CHOICE OF TAP LOCATIONS

VIEWGRAPH #15

B-282-89

My summary comments are that frequency domain techniques look really interesting and have a lot of potential and in some applications, where the adaptation can be internalized, already work wonderfully. But for all these other things we'd like to be able to do with it in a communications situation, particularly decision direction and blind equalizers that are nonlinear in some sense, there's a lot of work to be done.

HALL: John's up at Stanford and he did some of the original work on the fast Kalman techniques, and has promised us a look at multitone equalizers.

JOHN CIOFFI: *Algorithms for Multipath, Multitone Equalizers*

Actually the title of the talk listed two things: multipath fading and multitone equalization. [FOIL #1] In putting together the talk, I decided I'd stick with just one of those two topics, and I'm not going to say anything about the tracking of multipath fades today. I'd prefer to focus on an area that's kind of grown pretty quickly and rapidly in the last couple of years, the area of combining equalization and coding. Let me just give you an outline slide. [FOIL #2] We're looking at channels with intersymbol interference, and we're trying to design codes for those types of channels. So basically we're trying to combine both good distance properties and spectral shaping properties in a single entity. So I'd like to begin by just looking at some capacity arguments for intersymbol interference channels, and then looking at the variety of different types of gains that you can get, and then finally focus on the multichannel methods as a means to provide the best possible gains that you could achieve.

Now this is perhaps close to Figure 1 in any modern communications textbook. [FOIL #3] The key thing to note here is this is not an additive white gaussian noise channel, it's an additive gaussian noise channel – and we have a linear channel with some intersymbol interference, and the worse the ISI gets, the more interesting this particular problem becomes. Let me then couch that channel in a little broader setting [FOIL #4] where you can see the channel right here in the middle, and then we have some kind of transmit filter, some type of receive filter (which we call an equalizer in this particular case), and then outside of that we'd like to put on some good codes. Now here I'm referring to coset codes or trellis codes, which have arisen in the last 7 or 8 years as introduced by Ungerboeck and modified by many others. We'd like to concatenate those with the system to improve the overall gain for the system in terms of detection distance-to-noise ratio. The question here is, "Can you do any better?" The answer is "yes" if you have intersymbol interference in the channel and you combine the functions; that is, if you design the code and the filtering function together. So this extends the original Ungerboeck ideas where he combined both coding with just the modulation device, but basically did that for a channel which was free of intersymbol interference. How do you do it on a channel which has intersymbol interference? As I said, this problem has arisen in just the last couple of years, and there have been a number of solutions proposed to it in a fairly rapid period of time; I know Vedat Eyuboglu from Codex is going to be talking about another method of solving this problem this afternoon.

The classic result – and you can trace this all the way back to Shannon, but most people find it in Gallager's book on *Information Theory* – is what is the capacity of a channel that has intersymbol interference with additive gaussian noise like this? The classic

Multichannel Methods for Combined Equalization and Coding

May 16, 1989

John M. Cioffi

Information Systems Laboratory
Stanford University

FOIL #1

Outline

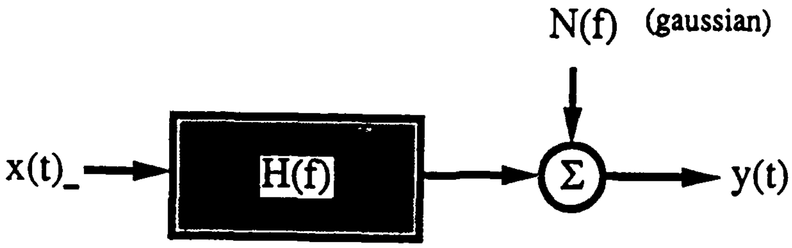
1. Capacity Arguments for ISI Channels

2. Separability of Gains
   - coding, shaping, equalization

3. Multichannel Methods

FOIL #2

Channel Model

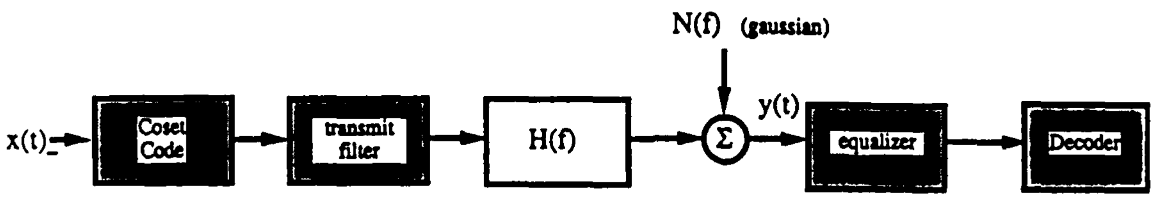


Linear channel with ISI

Additive gaussian noise

FOIL #3

Combined Equalization and Coding



Joint Design of the Equalization (Filtering) and Coding Functions

FOIL #4

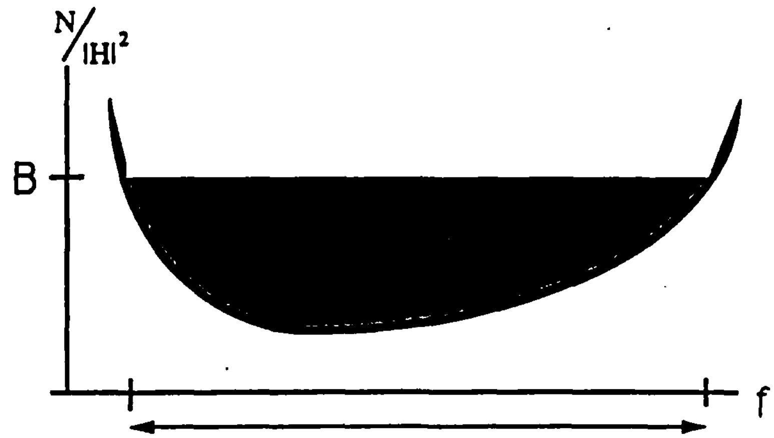
approach to computing that capacity is the so-called water filling approach. [FOIL #5] What you do is you take the channel transfer characteristic, in particular the power transfer characteristic, and you basically invert it, and then you take the amount of energy that you have to transmit, you have some kind of transmit power constraint, and you fill this bowl here, and that's what they call "water pouring". You pour water or energy into this bowl-type shape until you get to a level where you have basically the total of all the energy in there is equal to the power constraint that you have in the transmitter. Associated with that is some optimum bandwidth that you can use on the particular channel. Then what the capacity calculation basically does is it takes the capacity at each of the infinitesimally small channels in the frequency domain and sums them, via an integral, over the entire range, and that will be the capacity for the overall system with intersymbol interference. So what I'd like to do is take an example and look at this capacity calculation, and compare it with what the conventional approach would be on a channel with intersymbol interference in this particular situation.

So what we have here are four curves. The rightmost curve, or the best curve of the bunch [FOIL #6], is an additive white gaussian noise channel which you can compute using a well known formula: $1/2 \log(1+\mathrm{SNR})$ is the capacity for that channel. I've plotted it, the vertical axis is bits/symbol or bits/T; and the horizontal axis here is the channel signal-to-noise ratio. Now one of my students, Glen Dudevoir, has also done it for a fairly simple type of channel, it's just got one intersymbol interference term, $1 + 0.9D$, so that means you have a center tap on the channel and then delayed one unit in time you have another sample which is 90% of the amplitude of the first sample. I've looked at what the capacity of that particular channel is using the water-filling technique, and that is the dark line, which is somewhat less than the original channel, but basically parallels it, especially at the high signal-to-noise ratios.

Now this particular curve, right here, the third curve, is basically the performance that you get using pulse amplitude modulation. This is a single dimensional channel we're looking at here. What the performance would be if you had a certain signal-to-noise ratio, say 12 dB, and you wanted to achieve a certain error rate (10^{-6} is what we used here), how many bits/symbol could you get over the channel with that particular error rate? There's a fairly well known result, due to Dave Forney, that the flat PAM case is basically 9 dB worse at error rate 10^{-6} in the channel capacity, but basically it tracks that curve just to the right of it, and they're displaced this 9 dB difference. Now if you use very good trellis codes, using good codes with good shaping, you can basically achieve about – these days, the best I know of is about 7.1 dB of coding gain – so that you're about 1.7 dB away from that capacity curve. That's using the 256-state trellis code combined with some of the very recent trellis shaping methods. You can basically get that close to the curve, at least at high signal-to-noise ratio. So basically you're almost there at capacity on the flat channel.

Now what happens if you put a linear feedforward equalizer, the type that John Proakis described earlier this morning, on the $1+0.9D$ channel, you try to equalize it to be flat, and then look at the performance of that system in much the same way that we've looked at the performance of the flat channel. In other words, we look at error rate 10^{-6} and see what

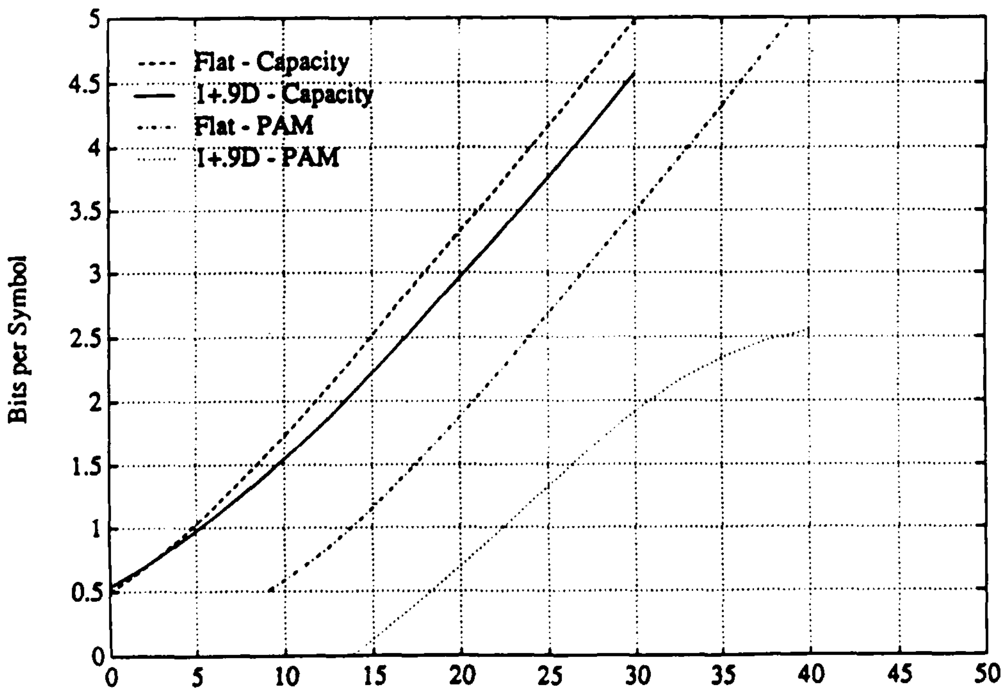
Capacity Calculation by Water-Pouring



optimum bandwidth Ω

$$C = \frac{1}{2}\int_{\Omega} \log\left(\frac{B|H(f)|^2}{N(f)}\right) df$$

FOIL #5



FOIL #6

the data rate that could be achieved for any signal-to-noise ratio is. And you see it's more than 9 dB away from its corresponding capacity curve. In fact, as you get out to very high signal-to-noise ratios it falls off considerably. So this is a motivation, just slapping a feedforward equalizer on and concatenating it with the code is not going to work unless the channel is pretty close to flat, and the more severe the ISI on the channel, the further off it's going to be. So that's really not the way to try to combine your codes with equalization on this particular channel.

Now I've also – one of my students, Jim Aslanis, did a couple of other plots of this same sort for two different channels. [FOIL #7] One is the classic $1 - D$ channel, AMI channel if you like, up on the top plot, here. Basically we're looking at, in this case, the performance of a decision feedback equalizer rather than a feedforward equalizer. The solid curve is, again, the capacity as computed via the water filling type plots. The second curve is using a decision feedback equalizer and concatenating it with the best trellis code you can find. Look at the difference between these two curves. Now for the $1 - D$ channel you can see the decision feedback equalizer performs a little bit better, it's closer to that capacity curve, and it basically hugs it at high signal-to-noise ratios. But as the intersymbol interference gets more severe on the channel, as in the bottom curve (I've got a channel where I've thrown several additional $1 + D$ factors at it, and what that does is force the channel spectrum towards low frequencies), this gap widens. The decision feedback approach is not so good on this particular channel, but again you can see these two curves merging as you go off to very high signal-to-noise ratio. This is a classic result that's basically due to Price twenty years ago, where he showed that the decision feedback equalizer is about as good as you can ever expect to do on a channel at high signal-to-noise ratio. Well, high signal-to-noise ratio may be extremely high in certain practical applications, and the more severe your ISI is on the channel, the higher that high SNR has to be before that argument basically holds true. So we have this motivation: present systems using the typical type of equalization on a channel with severe ISI really doesn't get where capacity arguments would tell you you should be on the channel.

So what I'd like to do now is look at how you might get there [FOIL #8], and look at, at least, some fundamentals which you can identify in that particular process. So I've returned here to the slide I began with [FOIL #9], describing combined equalization and coding, and I've augmented it by this formula right here. There are three terms. Basically γ is the overall coding gain for the system; that's the ratio of minimum distance in your coded system to the energy that it requires to transmit with that given minimum distance, and then you compare that for a well coded system with a system which is basically pulse amplitude modulation and uncoded. And there are three terms that you can identify in terms of the performance of this system. Two of them are fairly well known at this point – I think Dave Forney is the one who developed them in his coset codes papers, which just appeared last year, so-called "fundamental gain," which is basically a property of the trellis code itself without looking at the constellation shaping effects. That can go up to about 6 dB on the best known codes. There's a second term called "shaping gain" which has to do solely with the actual shape of the constellation and how close to a sphere it is in N dimensions, and that can go up to about

Capacity for PR Channels
(Kasturia, Aslanis)

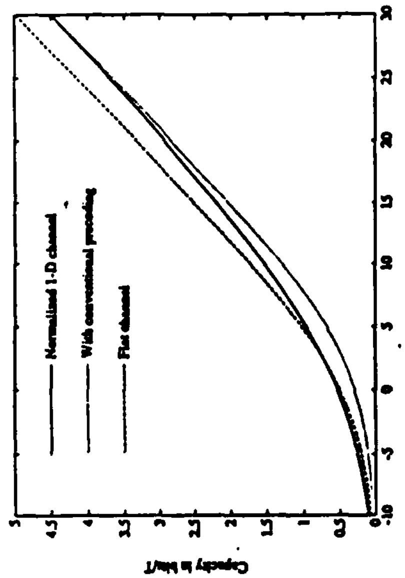


Figure 2: Capacity: normalized $1 - D$

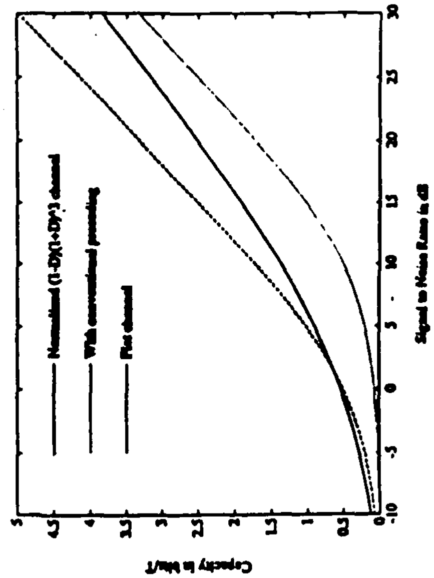


Figure 3: Capacity: normalized $(1 - D)(1 + D)^3$

FOIL #7

Outline

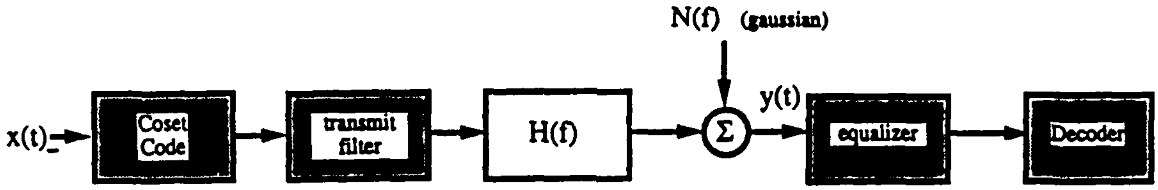
1. Capacity Arguments for ISI Channels

2. Separability of Gains
   - coding, shaping, **equalization**

3. Multichannel Methods

FOIL #8

Combined Equalization and Coding



$$\gamma = \gamma_f + \gamma_s + \gamma_e$$

fundamental gain $0 \le \gamma_f \le 6\text{dB}$

shaping gain $0 \le \gamma_s \le 1.1\text{ dB}$

equalization gain $0 \le \gamma_e \le$ **large !**

FOIL #9

Equalization Gain

$$\gamma_e = \frac{SNR_{MS\text{-}DFE,\Omega}}{SNR_{ZF\text{-}DFE,\ Nyquist}} = \frac{\widetilde{g_0^2} - \frac{1}{\overline{snr}}}{g_0^2}$$

g_0 = first tap in min-phase equiv. echannel

$\tilde{g}_0$ = tap 0 of MSPR channel for w. pour xmit filter

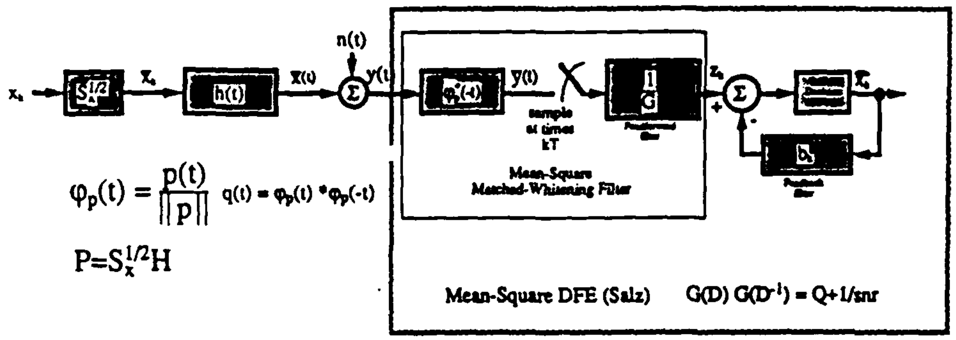
$$snr = \frac{\overline{E_X}\,|S_X^{1/2} H|^2}{\sigma^2}$$

FOIL #10

1.1 dB with the best known techniques for that. Then, finally, there's a third term that I'm going to add, and this is the "equalization gain." [FOIL #10] This is going to be a function of the channel that you're actually trying to use. On a flat channel, or a channel with very low signal-to-noise ratio, you basically have 0 dB of equalization gain that you can achieve. But on some channels this number can be very large, and so it's important to look at this and what it could be, and then how would we achieve it. So that third term, the equalization gain, I have defined this way: it's the signal-to-noise ratio of what I call the mean square error decision feedback equalizer; this is Salz's decision feedback equalizer, if you're familiar with that particular result. Basically it uses the least squares techniques rather than a zero forcing technique to design the decision feedback equalizer. And I use, for that equalizer, that Ω – that's the bandwidth that we saw earlier from the water pouring calculation. So you find that optimum energy distribution, use that with the decision feedback equalizer, and that's what I'm going to say is about the best that you're going to be able to do on your system. Then I compare that against what you might nominally try in a system like this, a zero forcing decision feedback equalizer, where you use the entire Nyquist band – you just put ± 1's or an i.i.d. data sequence, if you will, into the input of the channel. This is what I'll call the equalization gain. When you define it this way it basically adds in that third term in the formula you saw on the previous viewgraph. You can actually prove – I don't want to do it here – but it follows this simple formula here, where g_0 is the first tap in a so-called minimum phase equivalent for whatever the ISI channel is; $\bar{g}_0$ is the first tap or tap zero of what I call the mean square partial response channel, or the water pouring energy distribution ... I'll talk a little bit more about that. Basically what it is, is a least squares equivalent of this minimum phase channel. Then, finally, the signal-to-noise ratio that I use here, the little snr (using lower case letters) is just basically the $\frac{\text{transmit energy} \times \text{filter response}}{\text{noise}}$, which is a constant for any particular channel. So that's the equalization gain.

I promised I'd tell you a little bit more about what that mean square partial response, or least squares minimum phase channel, is. [FOIL #11] What it is, is if you take your data sequence, you run it through some kind of transmit filter where this transmit filter basically has the shape, a square root of the shape, which is associated with the water pouring energy distribution. You run through the channel. What you use over here – this looks like a decision feedback equalizer, it's actually Salz's decision feedback equalizer which is based on a mean square error criteria – if you run that through the feedforward section of his decision feedback equalizer, what will happen here is that you don't quite get the minimum phase equivalent for the channel. You get, what I call, a mean square error minimum phase channel. That was what that $\bar{g}_0$ was associated with; it's the lead tap in this least squares equivalent channel. If you want a formula for that, you just take the channel autocorrelation function, normalize it, and add $\frac{1}{\text{channel SNR}}$ to it, and then do a spectral factorization for that. This g here is the quantity that you're interested in. There are cepstral formulas that you can work out for this thing as well. The most interesting result that I found, about a year ago, was that if you look at the signal-to-noise ratio for this particular decision feedback equalizer, the minimum mean square error decision

Mean Square Partial Response



$$\varphi_p(t) = \frac{p(t)}{\|p\|} \quad q(t) = \varphi_p(t) * \varphi_p(-t)$$

$$P = S_x^{1/2} H$$

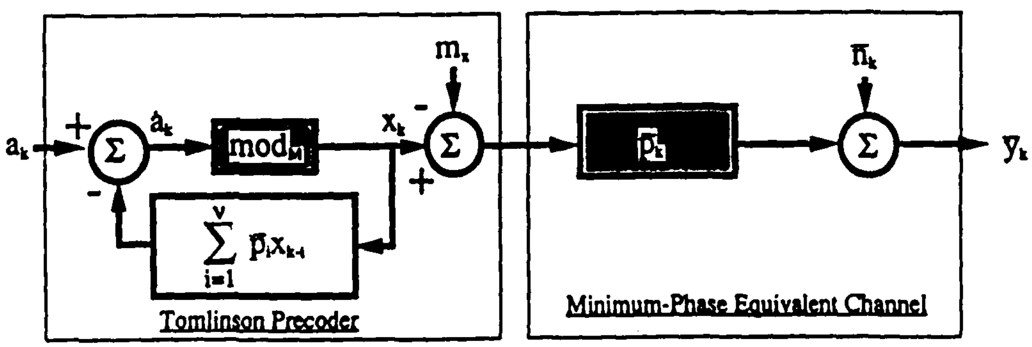
$$G(D)\,G(D^{-1}) = Q + 1/snr$$

$$g_0^2 = \int_{-1/2}^{1/2} \ln(Q)\, df \qquad \bar{g}_0^2 = \int_{-1/2}^{1/2} \ln(Q + 1/snr)\, df$$

$$C = \frac{1}{2} \log_2(1 + SNR_{MS\text{-}DFE,\Omega})$$

FOIL #11

Tomlinson Precoder



Eliminates Error Propagation

loss of shaping, xmit power boost

FOIL #12

feedback equalizer, when you use the water filling bandwidth, it turns out that the capacity for the channel with intersymbol interference follows this particular formula here: it's $1/2 \log_2 (1+\text{SNR})$ for this particular system. Now that is a very convenient form, because what you can see is that this formula is very similar to the thing that you get for a flat channel except that you just have the signal-to-noise ratio replaced by this mean square error decision feedback equalizer result. So what that says is now if you know how to design a code, 7.1 dB of gain, for a channel which is flat, which is effectively what the performance of this decision feedback channel is, you can basically do as well as you could do on the flat channel, just using a signal-to-noise ratio in place of that.

A classic problem with decision feedback is error propagation, which was not included in that last result. That's a very significant problem in a system that uses a decision feedback equalizer when you try to concatenate that with a code, because the error rate on the DFE output is very poor. If you use a very good code it's going to be about 9 dB (6 dB code + 3 dB signal expansion) away from what the actual coded performance is if you're just doing symbol-by-symbol detection, and that's an enormous problem in those types of systems. So the recent approaches to this have been to use the so-called Tomlinson precoder [FOIL #12] which basically shifts the feedback section to the transmitter with a very small penalty in transmit power boost, and some small penalty in a loss of shaping gain as well. So that's just a way of getting around that particular difficulty.

If you look at that particular system, these are the types of gains that you see for a variety of different partial response polynomials [FOIL #13], as were computed by Sanjay Kasturia, a past advisee of mine. The first channel, the $1 - D$, the gain is largest at lower bits/T, and it basically goes to zero as you get out to higher bits/T. As you have more severe ISI, there are huge gains to be had. Now some of these channels are a little bit ridiculous in that they have very low frequency content and a high Nyquist rate. But basically what's going on here is optimizing the transmit bandwidth using these techniques is a very important result that can show up in your overall gains. Now that's a basic concept of that the equalization gain – [I know I'm running out of time here, I've been getting some fingers, two and three, shown to me – all polite ones, right?]

The approach that I've looked at most for this is so-called multitone or multichannel type signaling techniques [FOIL #14]. Basically what you do is you split your channel up a la the same types of techniques that John Treichler was talking about [FOIL #15], with the transmultiplexing techniques, into a variety of parallel channels, and you try to again allocate your energy from the transmitter according to some water pouring assignment – it turns out that it's the same one that we saw earlier. When you do that and you fix the signal-to-noise ratio on all the outputs to your system to be constant, then maximize that signal-to-noise ratio, it turns out that this thing we saw earlier for the mean square error decision feedback equalizer over that water pouring bandwidth is the same as the signal-to-noise ratio that you get on the multitone system when optimized over that same bandwidth. And in fact you get this same capacity result where you could stick in either of the signal-to-noise ratios there. So this is another technique for doing it. It has no feedback in it and you basically go to these parallel channel type approaches.

MSPR Gains γ_e (dB)

	SNR 1bit/T	2bit/T	3bit/T	4bit/T
(1-D)	1.5	.8	.3	.1
$(1+D)^2(1-D)$	4.2	2.8	1.7	1.0
$(1+D)^3(1-D)$	8	6	4	2.9
$(1+D)^4(1-D)$	12.5	9.5	7	5

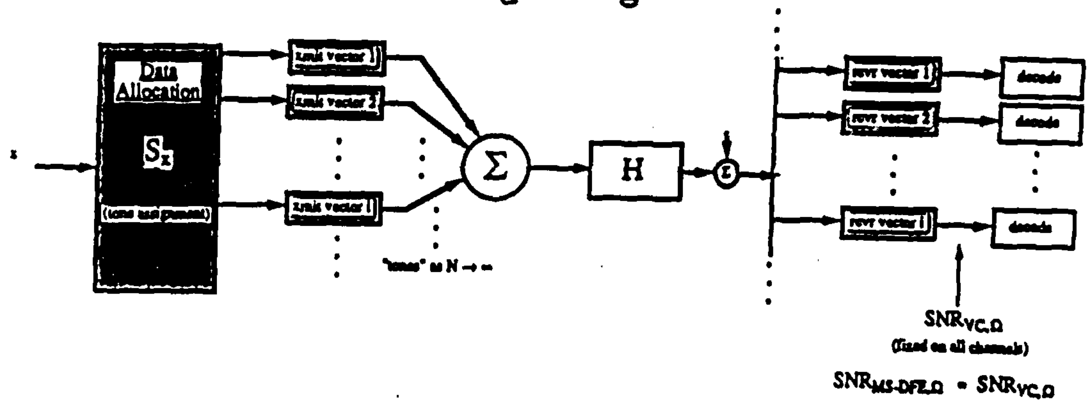
FOIL #13

Outline

1. Capacity Arguments for ISI Channels
2. Separability of Gains
   - coding, shaping, <u>equalization</u>
3. Multichannel Methods

FOIL #14

Mutlichannel Signalling

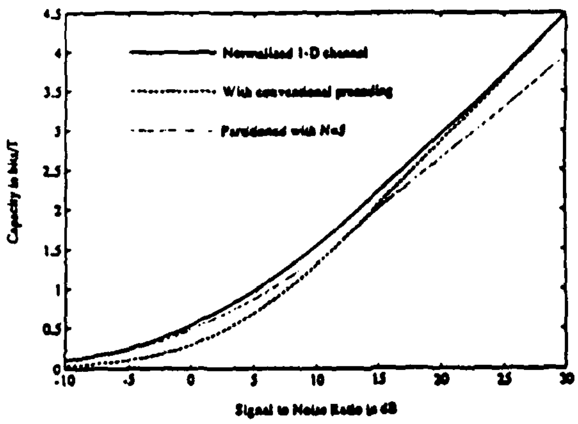


Infinite Dimensional Blocks

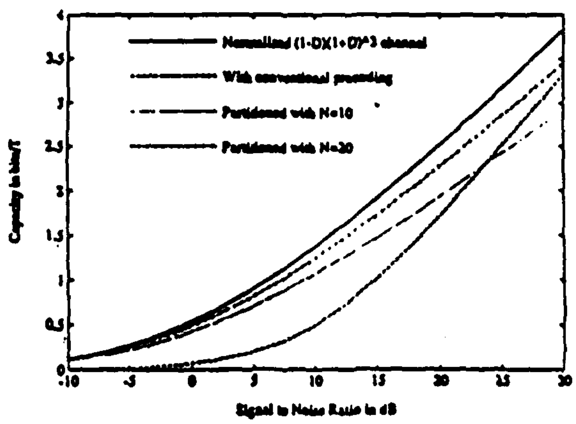
$$C = \frac{1}{2}\log_2(1+SNR_{MS\text{-}DFE,\Omega})$$

FOIL #15

Capacity - parallel Channels



(1–D)



$(1+D)^3(1\text{-}D)$

FOIL #16

Briefly, then, if you return to capacity for these types of systems [FOIL #16] and then look at how many tones that you actually need in doing that, if you look on this particular curve which I had put up earlier and I showed you the big difference between the actual capacity and what you got using a decision feedback equalizer operating over the entire bandwidth, if you use a multitone type system with block length of 10, you can see there's some improvement in this range. Use block length of 20, which means 20 - block length of 20 means 10 tones, you see an improvement, which is greater and as you increase the number of tones, you'll basically hug the capacity plot, basically because you're using the optimum spectrum at the transmitter bandwidth.

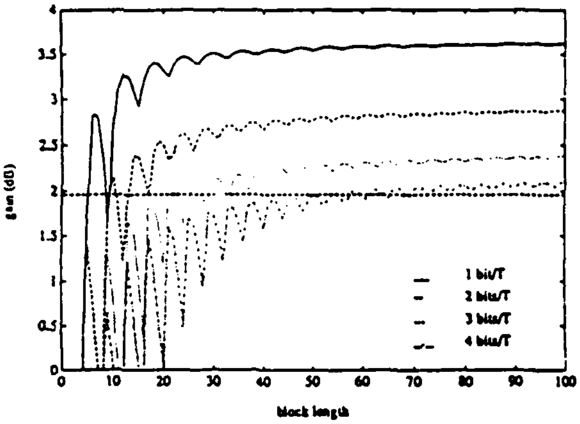
Now if you actually design a coded system to do that, these are some of the gains that you would see on the channel with respect to a decision feedback equalizer [FOIL #17] or, equivalently, the Tomlinson precoded system, where you use the entire bandwidth of the system rather than using the optimum transmit bandwidth. For the $1-D$ channel, again, you can see the same type of gain that you saw earlier for the partial response system. You see, this curve here - this dotted line across - is the decision feedback system. [FOIL #18] Up here you see the multitone-like system at 1 bit/T, 2 bits/T. And what we have here on the horizontal axis is the blocklength that you need, or basically twice the number of tones that you need in a multitone-like system to do it. You can see that the gain is fairly small for the $1-D$ channel, and then starts getting bigger as the intersymbol interference becomes more severe. You can throw up another plot where you have channels with yet more severe ISI, and basically you're seeing this same type of gain if you use the right number of tones in the overall system.

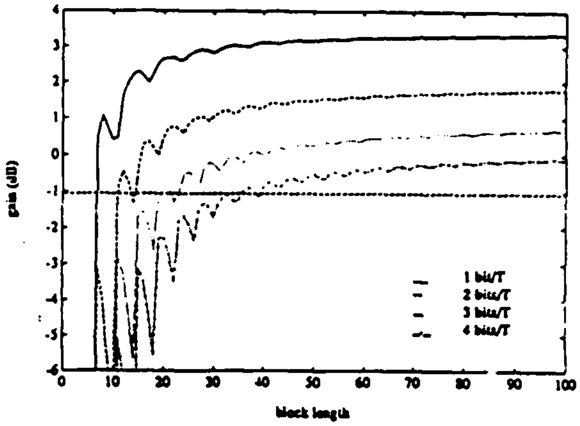
So I think, given that we're running out of time here, I'm going to skip the next couple of slides [FOILS #19,20] and just give you an overview of what's going on. This has been recognized in the last few years by a number of people - in fact, Telebit is now marketing voiceband modems [FOIL #21] which use the multitone-type schemes and use approximately the right bandwidth according to that, and they get very high speed. In fact they get 19.2 kilobits/second over voice grade wires using those types of techniques. It's also been recently proposed to the T1-E1.4 of ANSI for the so-called high rate ISDN service [FOIL #22], 800 kilobits/second or higher on twisted pair of loops within a carrier serving area. So it's starting to catch on as a technique to combine equalization and coding.

The scalar feedback things which I showed you will also basically get you to the same place. Using feedback has not - to the best of my knowledge - been used anywhere at present or contemplated for immediate use, but may well in fact wind up being used at some point in the future. This is a plot on that subscriber loop, that I was talking about versus input power and the megabits/second that you could put over. This was 9 kilofeet of 26-gage wire that we're looking at here. We have two different decision feedback alternatives, again using the standard approach using the entire transmitted bandwidth for the system at two different symbol rates, 400 and 800 kilosymbols/second. Then using basically what is a multitone-type system you can see that there's an enormous gain on this particular channel and that's why the interest is in these types of techniques in this subscriber loop.

Open topics - is there a convolutional type

Coding Gain

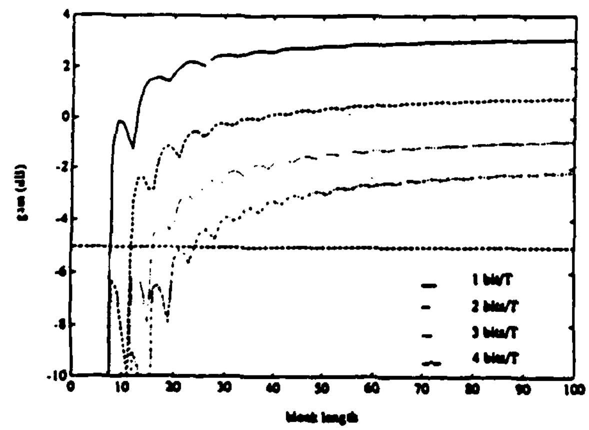


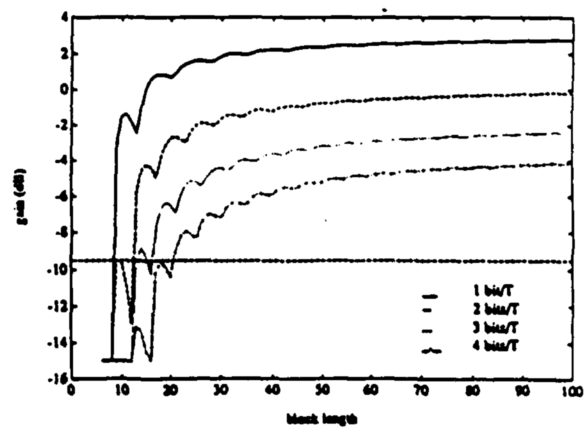


$1 - D$ $(1+D)^2(1-D)$

FOIL #17

Coding Gain

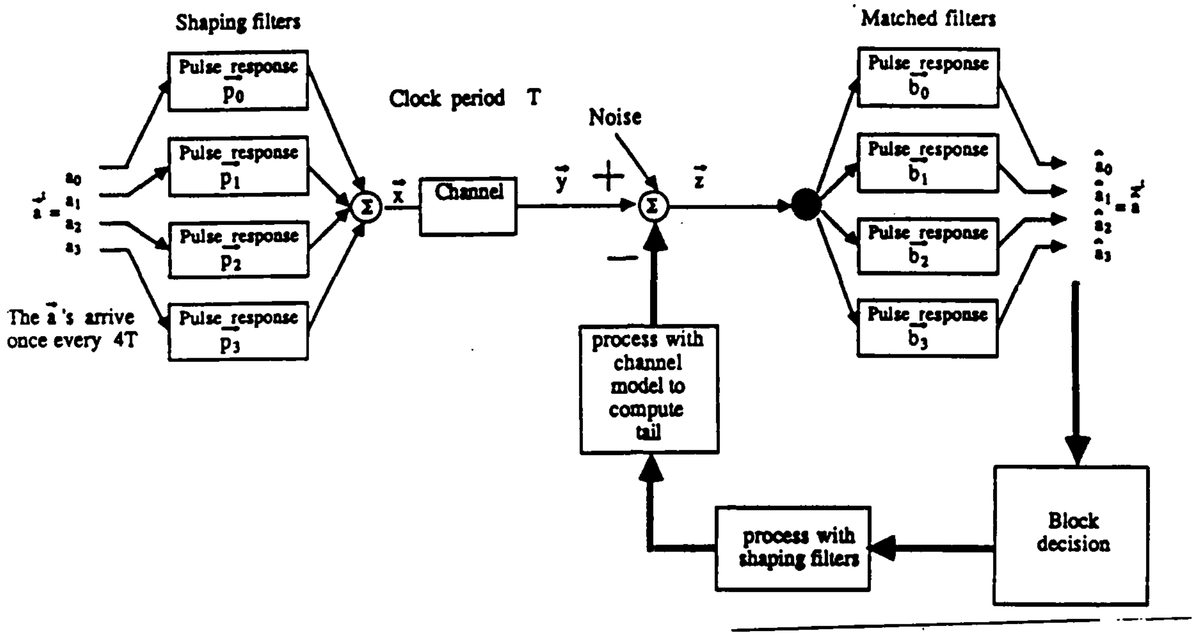




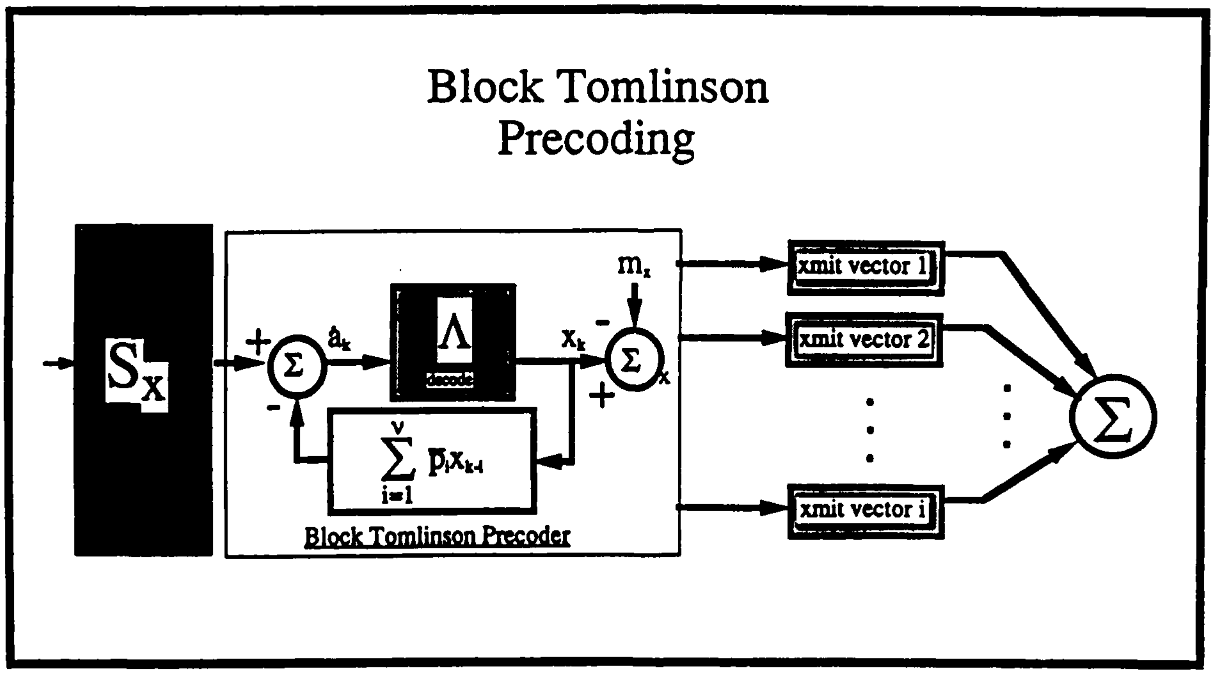
$(1+D)^3(1-D)$ $(1+D)^4(1-D)$

FOIL #18

BLOCK DFE (Kasturia) (Also, J. Lechleider , Bellcore)



FOIL #19



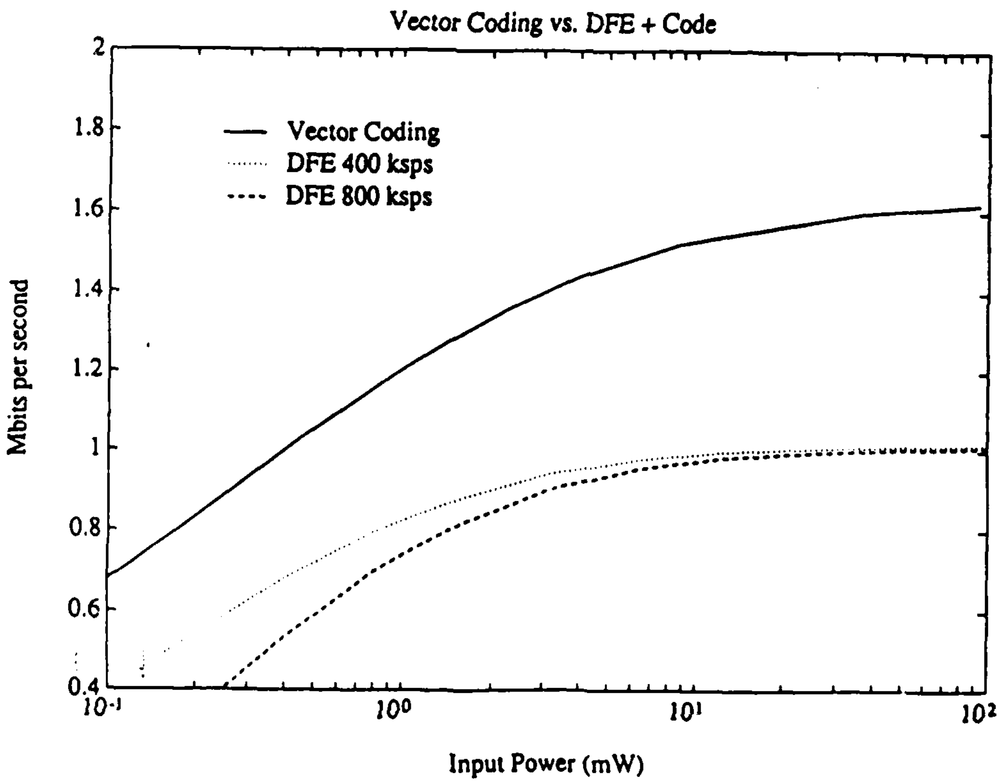
FOIL #20

Status of CEC

multitone - Telebit voiceband
modems, High-Rate ISDN

scalar feedback - ?

FOIL #21



FOIL #22

Open Topics

Convolutional Multitone ?

Interblock Coding

Sensitivity to Channel Knowledge

Parallelism

Crosstalking Channels

FOIL #23

Conclusions

two general methods for highest performance
multitone, mspr

can come within 1.7dB of capacity, at least with multitone at High SNR and within 3dB at low SNR

Application areas:

modems, T1 distribution on copper, copper LANs, and storage systems

FOIL #24

of multitone system where you could reduce the complexity of the system somewhat? Interblock coding – this means designing codes which span basically all the tones in the system rather than putting a separate trellis code on each tone in the system. How sensitive is this to actually knowing what your channel is at the transmitter? [FOIL #23] There's a significant amount of parallelism in the actual construction of a multitone system, and some of the problems you see, especially in the subscriber loop and the crosstalking channels area, those haven't been completely resolved as yet. It may be possible to improve those results I've shown you on the previous slide if someone can solve them.

So in conclusion, there are basically two methods that have arisen in the last couple of years, really, that people have proved are going to do as well on any channel with ISI as you can do on a flat channel if you use the right technique. The two techniques that are known to work right now are basically an optimized multitone technique or what I call mean square partial response techniques. Either of them will get you to the same performance level, and that performance level is within 1.7 dB of capacity at high signal-to-noise ratio, or if you have a really low signal-to-noise ratio it's about 3 dB away. [FOIL #24] It's already been applied in voiceband modems that are being marketed, and suggested for T1 distribution, this is 1.5 megabits on copper twisted pairs. There has been some interest I saw at ATT in looking at copper local area networks using these types of techniques. Most recently people have been looking at these techniques for storage systems as well. So that basically concludes the talk ....

HALL: Thank you very much. Our final speaker is Brian Agee, who co-published a paper several years ago with John Treichler on the constant modulus, and now he's got his own company and he's pursuing similar things ....

BRIAN AGEE: *Blind Adaptive Signal Restoration*

Thank you. I am going to be talking today about a general approach for designing structures and algorithms for blind adaptive signal extraction, which I refer to as the *property restoral approach* to blind adaptive signal extraction. I believe that this approach is very pertinent to the topics that we have been discussing at this workshop, especially the topics that were discussed at the yesterday's morning session. In particular, I think that this technique specifically addresses the problem that was raised by Dr. Pickholtz at that session, of "How can we separate and sort cochannel signals in a given environment?" I should also mention that this talk is going to be a bit of a change of pace from the other three presentations in this session, in that I am not going to be talking about equalization *per se*, but rather *any* technique for extracting signals from environments containing interference and linear channel distortion.

Before I start talking about blind adaptation, let me very quickly review the conventional adaptation approach, to show where blind adaptation departs from it. Slide AGI-2 depicts a typical non-blind adaptive processor. A vector sequence $\mathbf{x}(n)$ is formed from some received data set, using some fixed preprocessing stage – for instance, by passing the received signal through a tap delay line filter and setting $\mathbf{x}(n)$ equal to the tap voltages, or by setting $\mathbf{x}(n)$ equal to the voltages on an array of sensors. This vector signal is then passed through a linear combiner to yield output signal setting $y(n)$, and the weights of the combiner are adapted to optimize some

AGI

THE PROPERTY-RESTORAL APPROACH TO BLIND ADAPTIVE SIGNAL EXTRACTION*

Brian G. Agee
AGI Engineering Consulting
Woodland, CA 95695

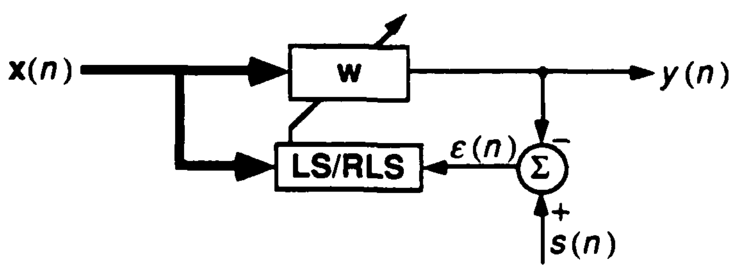
* Portions of the work presented here have been funded by ESL, Inc. (with partially matching funding from the California State MICRO Program), ARGOSystems, Inc., E-Systems, Inc. and the National Science Foundation

AGI-1

AGI

BACKGROUND: CONVENTIONAL ADAPTATION

•Example: MMSE linear processor



• Goal: adapt processor to restore SOI *quality*
  - minimize MSE $<|\varepsilon(n)|^2>$, BER, etc.
  - maximize SINR, likelihood function, etc.

•Problem: requires knowledge of SOI waveform or channel

AGI-2

measure of the *quality* of the signal of interest (SOI) $s(n)$ within $y(n)$ (in this case the mean-square-error between $y(n)$ and $s(n)$).

However, in order to perform this optimization, the receiver requires knowledge of either the SOI waveform or the transmission channel that the signal was transmitted over. In many applications, this information is not available at the receiver, and quality optimization cannot be performed. Instead, we must use *blind* adaptation techniques that can operate without the use of this information.

Slide AGI-3 lists the assumptions that we are operating under here that motivate the use of blind algorithms. We are going to assume that the SOI is unknown over the reception interval, so we cannot use it for adaptation. We are also going to assume that the SOI transmission channel is also unknown at the transmission start, and possibly time-varying over the reception interval. In an antenna array application, for instance, these assumptions would be consistent with applications where we do not have calibration data available at the receiver, or where our array geometry is changing too quickly for us to track changes in the cal data. In addition, we are going to assume that the channel itself may be time varying, for instance due to interferers coming up or going down over the reception interval.

So, the question is, "What can we do in this situation?" We can't restore the SOI quality, because we can't *observe* the SOI quality: we do not have access to either the transmitted signal waveform or transmission channel. However, in many applications we *do* know something about the SOI: it has some known structure or *properties*. What we can therefore do is try to optimize some observable measure of these known signal properties, rather than some unobservable measure of the signal quality. This leads to the *property restoral approach.*

In Slide AGI-5, I have listed a 4-step design methodology for designing a property restoral algorithm. This procedure is as follows.

Step 1: Identify the exploitable properties of the SOI. Analyze the structure of the SOI, and determine the properties of that signal that distinguish it from the background interference.

Step 2: Identify a signal extraction structure that can perform two functions: 1) *extract* the SOI from the received environment, that is, remove the interference and distortion from the received signal (the usual function); and 2) *restore* the particular SOI properties that we are trying to *exploit*. The second function adds a subtle constraint to the conventional procedure for designing an extraction processor: the extraction must be performed in a *nontrivial* manner that does not inadvertently impart the desired signal properties to the processor output signal. In particular, the processor must be constrained so that restoral of the signal properties is tantamount to restoral of the signal quality.

Step 3: Design an objective function that is optimized when the SOI properties identified in Step 1 have been imparted to the signal at the output of the processor designed in Step 2.

Step 4: Design an optimization algorithm to find the useful extrema of the objective function designed in Step 3. Again, the desired outcome is for restoral of the signal properties to be tantamount to

BLIND SIGNAL EXTRACTION PROBLEM

- Assumptions
  - SOI unknown over reception interval
  - SOI transmission channel unknown at reception start, possibly time-varying over reception interval
  - Overall channel possibly *dynamic* over reception interval
- Applications
  - ESM systems (reconnaissance, acquisition)
  - EW/ECM systems (anti-jam, look-through)
  - Mobile & satellite communications
  - Low-cost telephony, data communications
  - Reception of broadcast signals

AGI-3

HISTORICAL PERSPECTIVE

- Partially-blind techniques
  - Linearly-constrained BF (Griffiths '69, Frost '72)
  - Noise cancellation (Widrow '75)
  - Decision directed equalization (Proakis '69, Gersho '69)
  - Decision feedback equalization (Austin '67, Monsen '71)
- Truly-blind techniques
  - Sato's algorithm (Sato '75)
  - Reduced constellation algorithm (Godard '80)
  - Dispersion-directed algs. (Godard '80, Benveniste '80-'84)
  - Constant/known modulus algorithms (Treichler '83, '85)

AGI-4

AGI PROPERTY-RESTORAL CONCEPT

- Adapt processor to restore known SOI properties
  - Modulus of FSK, PSK, FM SOIs
  - Self-coherence of PCM, AM, FDM-FM SOIs

- Basic design methodology

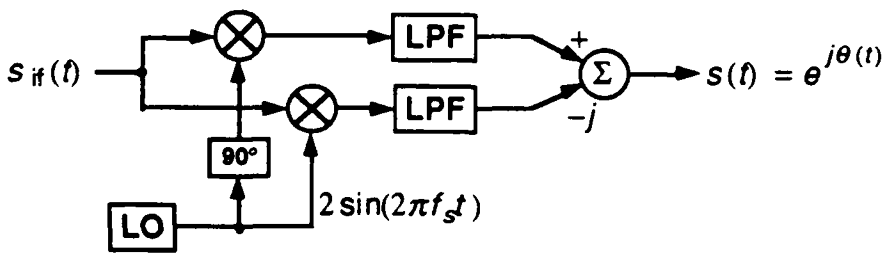
  1. Identify exploitable SOI properties
  2. Identify signal extractor structure
  3. Design property-restoral objective function
  4. Develop property-restoral adaptation algorithm

- Desired outcome: restored *property* ⇔ restored *quality*

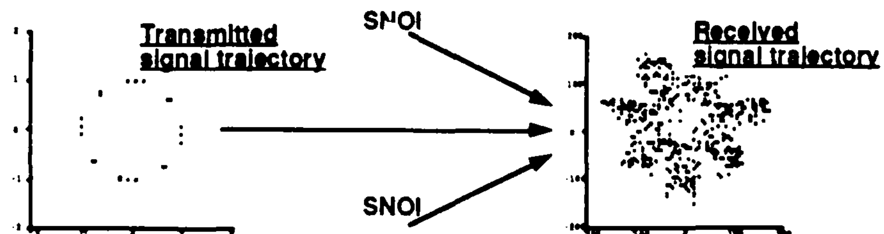
AGI-5

AGI EXAMPLE: CONSTANT MODULUS PROPERTY

•SOI possesses *constant complex envelope*



•Property destroyed by distortion, interference



AGI-6

restoral of the signal quality at the processor output.

Slides AGI-6 and AGI-7 illustrate this approach for the *constant modulus algorithm (CMA)*, which was developed by John Treichler and myself in the early 80's at ARGO Systems [1]. This was the first algorithm that (to my knowledge) explicitly used the property restoral concept in its development. At the time that we developed this algorithm, we were attempting to remove multipath interference from an FM signal. The basic approach was developed by noting that the FM signal can be represented as $s(t) = e^{j\theta(t)}$ after it is transformed to complex baseband representation. That means that the magnitude or *modulus* of the signal is *constant* (equal to unity). However, this "constant modulus" property is destroyed if interference or distortion is added to the FM signal. This motivates a simple modification of the nonblind processor shown in Slide AGI-2, by adapting the linear combiner to minimize a measure of the *modulus variation* of the signal at the processor output. The particular modulus variation measure shown in Slide AGI-7 is referred to here as the *1-2 constant modulus cost function*, which is my personal favorite for performance and implementation reasons. Essentially, this cost function is the mean-squared-error between the modulus of the processor output signal and unity. If $y(n)$ is equal to the transmitted signal, this cost function is equal to zero; if the modulus variation of $y(n)$ is low, the cost function value is also low. By adapting the linear combiner weights to minimize the cost function, we are hopefully going to extract the SOI from the received environment. In fact, this algorithm has worked very well in practice, and has been implemented (under various names) in both military and commercial communications systems.

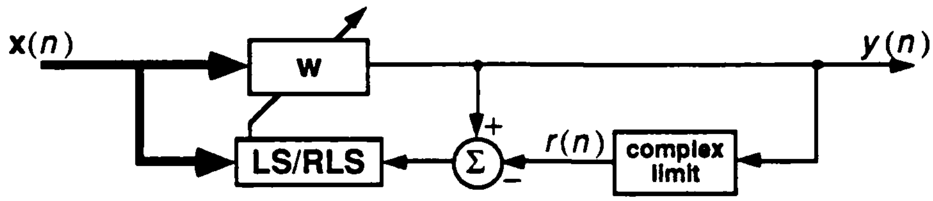
So, given this genesis, what's possible now? Slide AGI-8 lists what I think is possible now quite alot! The CMA was originally developed to adapt single-sensor FIR filters; since then, it has been applied to a number of different processor structures, including multisensor antenna arrays and polarization combiners [2], and it has been implemented in frequency-domain as well as time-domain form (as discussed in John Treichler's presentation). In addition, a number of fast and rapidly converging versions of the CMA have been developed by myself and others [3,4], which allows the blind processing concept to be applied to rapidly or dynamically time-varying environments. The property restoral concept has also been applied to a number of new processor structures, including adaptive *demodulators*, which restore properties of the *baseband symbol sequence* rather than the transmitted signal waveform [5,6], and *multitarget* processors, which extract *multiple* SOIs from the received environment [2,7]. This last extension addresses the sorting problem discussed in yesterday's morning session.

In addition, the modulus restoral approach has been applied to a number of other kinds of signal properties besides constant modulus. Most of these features fall into two general classes: modulus features, or features of the envelope of the signal; and self-coherence features or self-correlation of the SOI under temporal, spectral or spatial displacement. However, other features have also been exploited that fall into neither of these classes, or that represent generalizations of these classes. In future, I believe that it shall be possible to extend these approaches to a number of additional modulation features.

Lastly, we are starting to get a much stronger theoretical understanding of how

AGI CONSTANT MODULUS RESTORAL

•System Concept



•1-2 Constant-Modulus Cost Function

$$F_{1-2}(w) = \langle | y(n) - r(n)|^2 \rangle$$
$$= \langle (|y(n)| - 1)^2 \rangle$$

AGI-7

AGI WHAT'S POSSIBLE

- Blind extraction possible via many structures, algorithms
  - Antenna arrays, filters and demodulators
  - Single-target or multitarget (multiport) processors
  - Time-domain or frequency-domain implementations
  - Fast/rapidly-converging algorithms
- Can exploit a plethora of SOI modulation features
  - Modulus (envelope properties)
  - Spectral self-coherence (cyclostationarity)
  - General modulation properties (property-mapping)
- Strong theoretical understanding beginning to emerge
  - Uniqueness proofs (Benveniste, Godard)
  - General convergent behavior (Lundel, Agee)
  - Ties to maximum-likelihood estimation (Agee)
  - Ties to POCS, PM signal enhancement methods (Agee)

AGI-8

these algorithms work, from the original work performed in this area by Benveniste [8] and Godard [9], and from more recent work performed by John Lundel at Stanford and myself at AGI and U.C. Davis. In particular, I have mathematically proven stability and convergence of a general class of algorithms that I refer to as *property-mapping algorithms* [3,AGI6], which encompass the class of modulus restoral algorithms, and Lundel and I have independently shown that these stationary points can correspond to capture of the different signals in the environment under fairly representative conditions [6,10,11]. So, the stationary points of our objective function actually mean something in terms of SOI quality.

In my Doctoral research, I have also discovered some interesting ties between property restoral techniques and maximum likelihood estimators of the SOIs, and also ties to some of the signal enhancement methods that have been developed in the field of image restoration. In particular, the property-mapping approach that I have developed in [6] is closely related to the method of projection onto convex sets developed by Youla [12,13] and the property mapping signal enhancement method developed by Cadzow [14]. In fact, I believe that it should be possible to use Youla's and Cadzow's work to substantially generalize the blind processing techniques that have been developed to date.

Obviously, there is a lot more that I could talk about today than I have time for. Therefore, I'm going to focus on some specific results in a couple of the areas discussed above. In particular, I want to talk a little about the new kind of property restoral approaches that I and other people are starting to look at, and discuss some of the more interesting or novel structures for blind signal extraction that I have been working on lately. In particular, I am going to focus on those structures that are particularly applicable to the modulation characterization and sorting problems that people talked about in yesterday's morning session.

The modulation properties that to my knowledge have been looked at to date are shown on Slide AGI-9. As I said on the previous slide, these properties generally fall into two categories: modulus properties, and self-coherence properties. I have been looking primarily at modulus properties and *spectral* self-coherence properties [15], which are caused by cyclostationarity of the signal and commonly induced by periodic bauding, gating and mixing operations at the signal transmitter.

The modulus properties that I have investigated in my research are listed on Slide AGI-10. These properties are all extensions of the "constant modulus property," that I discussed earlier. The constant modulus property applies to signals with truly constant moduli, such as CPFSK signals and FM signals, as well as signals with low (but nonzero) modulus variation, such as AM signals (with low modulation index). The constant modulus property can also apply to signals whose *baseband* or symbol sequence has a low modulus variation, such as PSK and PCM QAM signals. The second class of properties, the "known modulus property," was originally investigated by Frost (see [2]), and is applicable to signals with a non-constant but nonrandom and known modulus, for instance pulse communication signals, pulse radar signals and return-to-zero FSK signals. To my knowledge, this is the second technique that was developed explicitly using the property restoral viewpoint.

The third class of properties, the "multiple

AGI EXPLOITABLE PROPERTIES

- Modulus properties
  - Constant modulus / low modulus-variation
  - Known modulus
  - Known modulus distribution
  - Almost-periodic modulus
- Self-coherence properties
  - Spatial self-coherence
  - Temporal self-coherence
  - Spectral self-coherence, conjugate self-coherence
  - Higher-order self-coherence
- Other SOI properties
  - Signal transience
  - General signal properties (ML blind signal estimation)

AGI-9

AGI MODULUS PROPERTIES

Property	Definition	Primary Application
Constant modulus	$\lvert s(t) \rvert = 1$	All processors
Known modulus	$\lvert s(t) \rvert = m(t)$	Filters
Multiple modulus	$\lvert s(t) \rvert \in \{m_l\}$	Demodulators
Almost-periodic modulus	$\lvert s(t) \rvert = \sum_k M_k e^{j2\pi\alpha_k t}$	Antenna arrays

AGI-10

modulus property," has been independently investigated by myself and several others; in particular Sethares, Rey and Johnson are reporting on a version at the 1989 ICASSP conference [16]. The multiple modulus property is applicable to SOIs with non-constant and random moduli, where the modulus distribution is defined over a discrete number of known levels. This property is held by the symbol sequence of OOK, ASK and PCM QAM (APK) signals.

The fourth class of modulus properties, the "almost-periodic modulus property," has been developed by myself to overcome some specific problems of the known modulus property. The almost-periodic modulus property is applicable to SOIs with unknown moduli that are almost-periodic at some known set of harmonics. This property is held by pulse communication signals, pulse radar signals and return-to-zero FSK signals. This property is also applicable to many kinds of cyclostationary signals, for instance PCM QAM signals.

Notice that there is a third column in this Table (on Slide AGI-10), addressing the primary application for these properties. As I said earlier, the processing structure that we choose to perform our signal extraction is very dependent on the properties that we're trying to exploit. Some of these properties are more applicable to certain processor structures than others. For instance, the known modulus property is not very applicable to memoryless antenna arrays, because we must know the *timing phase* as well as the shape of the SOI modulus, in order to adapt the array to restore the correct modulus. In most applications, however, the timing phase of the received SOI is not generally known. On the other hand, the known modulus property is highly exploitable in adaptive filters, because the filter can readily adjust the timing phase (as well as the shape) of the output signal modulus to that of our target modulus. Similarly, the multiple modulus algorithm is well suited to demodulator structures, which can restore properties of the baseband symbol sequence, but they're not very applicable to adaptive arrays or filters, which can only restore properties of the pre-demodulated signal, because the modulus of the signal *waveform* can vary wildly between baud centers even if it is restricted to a discrete number of levels at the center of each baud.

Slide AGI-11 describes the general approach that I've developed to exploit this modulus restoral, which I refer to here as the *property mapping approach.* Basically I define a *property mapping cost function* $F(\mathbf{w})$ as a distance measure between the processor output signal $y(t)$, and some SOI estimate $s(t)$. The processor output signal is assured to be in the linear subspace $\mathcal{L}_{\mathbf{x}}$ spanned by the received (and preprocessed) vector data $\mathbf{x}(t)$ while $s(t)$ is assumed to be a member of the *desired signal property set* $\mathcal{D}_s$, which is the set of all signals with the properties to be restored by the processor. The cost function is then minimized using an alternating projections approach, which converges monotonically to a stationary point of the cost function. For instance, if $\mathcal{D}_s$ is the set of all constant modulus waveforms (or symbol sequences), and the distance measure is the time-averaged squared error between $y(t)$ and $s(t)$, then $F(\mathbf{w})$ reduces to the 1-2 constant modulus cost function and the property mapping algorithm reduces to the least squares CMA described in [3]. Each of the other modulus properties shown on Slide AGI-10 can also be used to generate a property mapping cost function and minimizing algorithm using the same general approach.

THE PROPERTY-RESTORAL APPROACH TO BLIND ADAPTIVE SIGNAL EXTRACTION

PROPERTY-MAPPING APPROACH

- Alternating projections approach

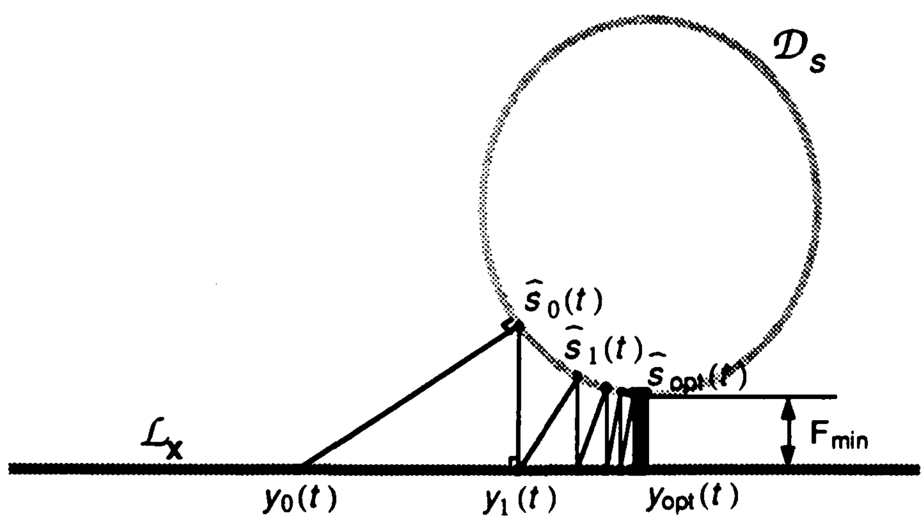
$$F_{min}(\mathbf{w}) \rightarrow d\left[y(t), \hat{s}(t)\right], \quad y(t) \in \mathcal{L}_X, \quad \hat{s}(t) \in \mathcal{D}_S$$

1. min F w.r.t. $\hat{s}(t)$, given $y(t)$
2. min F w.r.t. $y(t)$, given $\hat{s}(t)$

- Powerful convergence properties
  - Monotonic convergence to stationary point
  - Very benign capture properties
  - Much faster than steepest-descent
- Close ties to signal enhancement techniques
  - Projection-onto-convex-set (Youla, Trussell)
  - Property-mapping signal enhancement (Cadzow)
- Close tie to maximum-likelihood blind signal estimator

AGI-11

GEOMETRICAL INTERPRETATION



AGI-12

Slide AGI-12 provides a geometrical interpretation of the basic property mapping algorithm. Starting at some arbitrary data point in the data subspace $\mathcal{L}_{\mathbf{x}}$, the optimal value of $y(t)$ is arrived at by iteratively projecting y(t) onto the desired signal set $\mathcal{D}$, and then back onto the data space $\mathcal{L}_{\mathbf{x}}$. This process is continued until both $y(t)$ and $s(t)$ converge to a stationary solution; the final distance between these signals is then equal to the convergent cost function value. Anyone here who is familiar with the property mapping techniques developed by Youla or Cadzow should see some very strong similarities between these approaches. As I have said before, I think that this similarity can be exploited in the future to develop even more powerful property restoral algorithms.

In the next slides I want to discuss the structures that I have applied the property restoral approach to. The two particular structures that I will discuss today are multitarget antenna arrays, which I think will be of interest to people doing modulation characterization, and the adaptive linear PAM demodulators, which may also have eventual application to modulation characterization, in cases where sorting must be performed using single-sensor processors.

As the name indicates, the objective of the multitarget processor is to extract multiple signals from a received data set. The term "multitarget" is actually drawn from optimization theory, where algorithms for finding multiple stationary points or local extrema of an objective function are commonly referred to as multitarget optimization techniques. Multitarget algorithms can be used to not only extract multiple signals from the environment on the basis of their modulation properties, but also to sort the extracted signals and determine whether they are signals of interest or *not* of interest on the basis of those properties. Multitarget algorithms have been developed by myself for two classes of modulation properties: low modulus variation (constant modulus property), and known spectral self-coherence.

The multitarget property restoral algorithms developed to date by myself and others have primarily been applied to antenna array processors, because it is so much easier to separate signals with antenna arrays than with filters. I have also chosen to focus on antenna arrays because my understanding of the theoretical behavior of the blind adaptation algorithms that I have investigated to date is so much stronger for these kinds of processors. However, it should be possible to extend some of these algorithms to adaptive demodulators and filters as well.

Slides AGI-15 through AGI-18 present results of my work in developing a multitarget processor based on the least-squares CMA. I will be discussing this processor in more detail at the 1989 MILCOM conference [7]. The basic processor structure, shown in Slide AGI-15, is a multiport antenna array that uses a matrix beamformer to form a vector of output signals from the vector input signal. Each column or port of the beamformer is adapted in parallel using a least-squares CMA, with a soft orthogonality constraint applied to force each port to a different solution of the 1-2 constant modulus cost function, that is, to force the signals at the output of each port to have low correlation. In addition, algorithm modifications have been added to allow the processor to sort the output ports on the basis of modulus variation, in order to separate the active ports (which have captured signals) from the inactive ports (which have locked into the noise background); and to allow the processor to

AGI NEW STRUCTURES

- Multitarget processors (property-based sorting)
- Adaptive demodulators (baseband property-restoral)

AGI-13

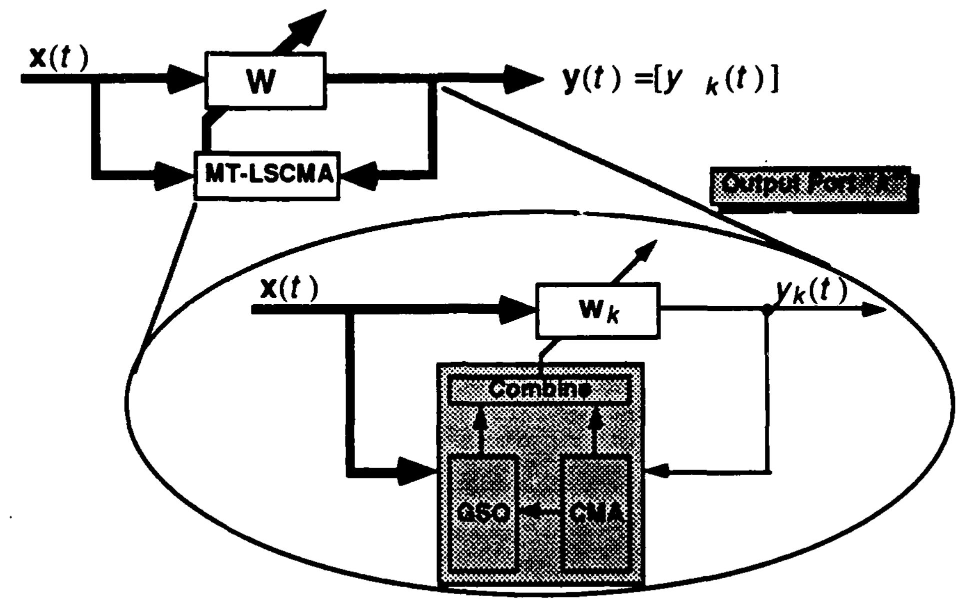
AGI MULTITARGET PROCESSORS

- Extract *multiple* signals from received environments
  - Primarily applicable to antenna array processors
  - Also applicable to adaptive demodulators
- Allow *sorting* on basis of modulation properties
  - Sorting via modulus properties
  - Sorting via self-coherence properties
- Structures developed to date
  - Multistage (sequential) structures (Treichler)
  - Multiport (parallel) structures (Agee)

AGI-14

AGI THE PROPERTY-RESTORAL APPROACH TO BLIND ADAPTIVE SIGNAL EXTRACTION

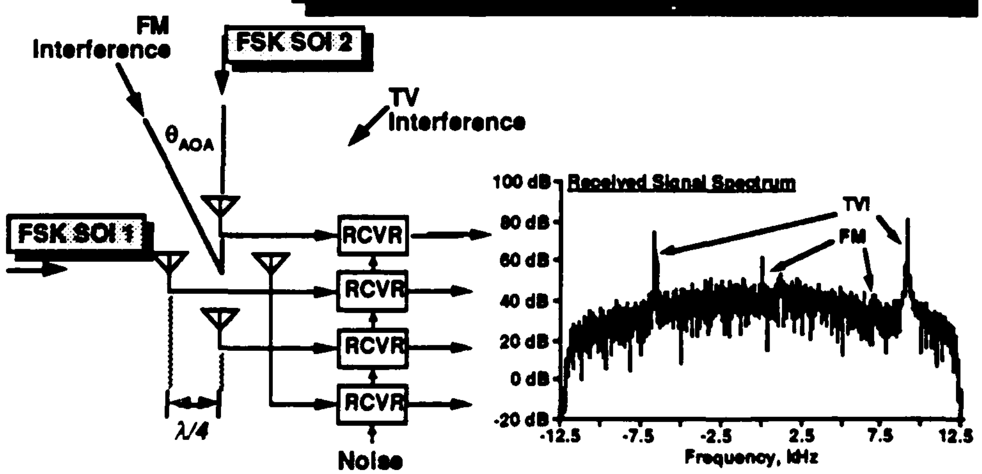
AGI EXAMPLE MULTIPORT STRUCTURE: MT-LSCMA



AGI-15

AGI BURST SIGNAL EXTRACTION EXAMPLE

Signal	Type	Carrier	ON Time	SWNR	AOA
0	FSK SOI 1	0 Hz	20 msec	17 dB	90°
1	FSK SOI 2	0 Hz	24 msec	20 dB	0°
2	NBFM	0 Hz	—	40 dB	30°
3	TVI	-5 kHz	—	50 dB	-45°



AGI-16

quickly capture signals as they appear in the environment and maintain capture on those signals as the environment varies.

This processor is illustrated in the next three slides (Slides AGI-16 through AGI-18) for a four-element narrowband antenna array excited by white thermal noise, two very strong interference signals – a TV signal at a 50 dB signal-to-white-noise-ratio (SWNR) and an FM signal at a 40 dB SWNR – and by two burst FSK SOIs. The FSK signals are received at a 17 dB SWNR and a 20 dB SWNR, respectively, which is very far below the interferers, and have ON and OFF times that cause them to *collide* over the reception interval: the first FSK signals comes ON at 20 msec into the collect and goes OFF at 30 msec into the collect, while the second FSK signal comes ON about 24 msec into the collect – right in the middle of the first FSK signal's ON time – and goes OFF at about 24 msec into the reception interval.

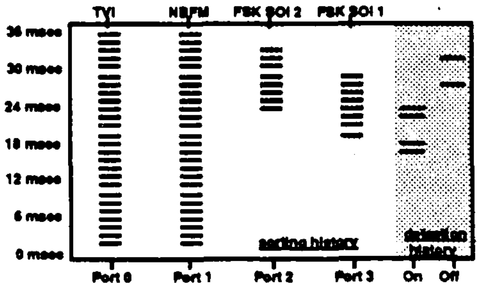
This is a very difficult kind of environment for any blind or nonblind adaptive processor to deal with. However, the multitarget LSCMA is able to collect both signals without bit error in this collect. Essentially, the algorithm converges to all the stationary points of the cost function. At the beginning of the collect, Port 0 and Port 1 capture the two interferers, and identify the remaining ports as "inactive." The active signal ports maintain capture on the interferers over the reception interval, while Ports 3 and 2 captures the first and second FSK signals, respectively, over their ON times. Note that Port 3 maintains capture on the first FSK signal, even when the other signal comes up, with some suboptimal performance (which it quickly recovers from) right when the second signal comes ON. Similarly, Port 2 maintains capture on the second FSK signal even when the first FSK signal disappears from the environment, without any appreciable deviation from its optimal (time varying) performance.

We can also use known spectral self-coherence properties of the SOIs to adapt a multitarget processor. An advantage of this approach over the CMA-adapted multitarget antenna array is that the SCORE processor is easily adapted to *exactly* optimize our objective function, and the algorithm extracts *only* the signals with the desired self-coherence properties from the received data. This is illustrated in Slides AGI-20 through AGI-23, for the *cross-SCORE* processor [15] shown on Slide AGI-19, and for a four-element antenna array excited by two very strong interferers, a BPSK SOI with a 4 MHz data rate, and a 16 QAM signal with a 3 MHz data rate. By tuning the target spectral self-coherence frequency a of the processor to 4 MHz and 3 MHz, the processor is able to extract each of the SOIs from the received data. This is accomplished *without* using knowledge of the direction of arrival of the SOIs, and without using any array calibration data either the array manifold or the background noise covariance.

The second class of processor structures that I want to discuss today is the *adaptive demodulator structures.* The demodulator structure assumed for linear-PAM SOIs is shown in Slide AGI-25; note that the demodulator reduces to a simple fractionally tapped equalizer (FSE) for this type of signal. The blind demodulator structure has two advantages over the equivalent adaptive filter: it attempts to restore the properties of the baseband *symbol sequence*, rather than the entire signal waveform; and it accomplishes this restoral using a processor that is much more powerful than an LTI filter. The first advantage obtains because the properties of the sig-

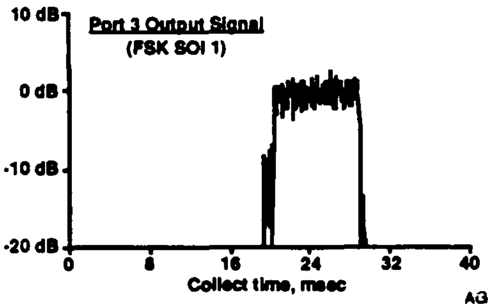
THE PROPERTY-RESTORAL APPROACH TO BLIND ADAPTIVE SIGNAL EXTRACTION

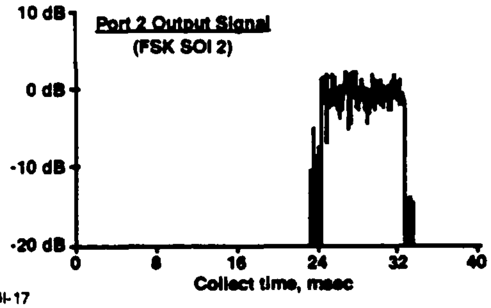
AGI ACQUISITION PERFORMANCE



Control Parameters

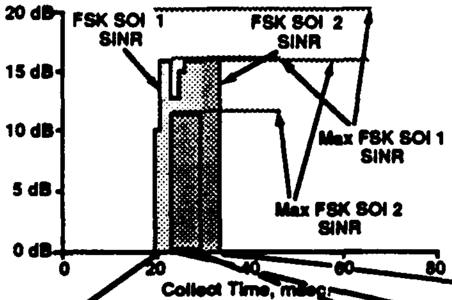
• Block size	1.28	msec
• Delay	2.56	msec
• Acq. threshold	7.00	dB
• Noncapture blanking	ON	



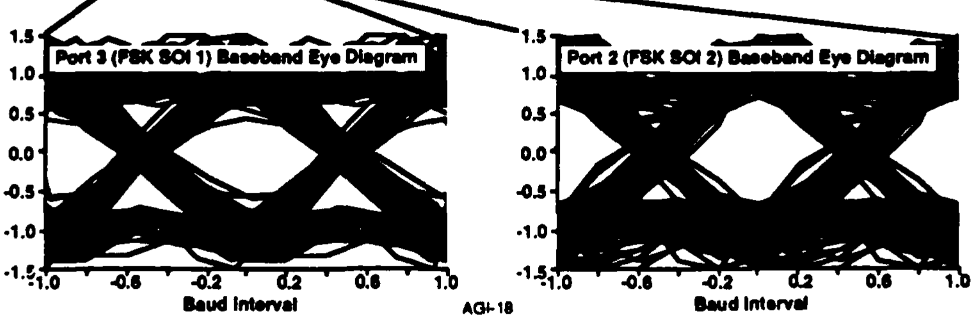


AGI-17

AGI EXTRACTION PERFORMANCE

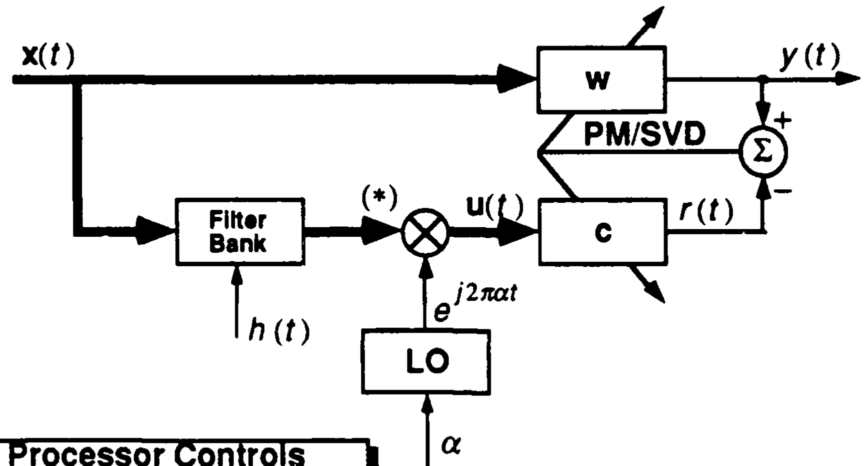


Port 3 Output	Port 2 Output
• No bit errors	• No bit errors
• 28% peak ISI	• 38% peak ISI
• 13% peak jitter	• 22% peak jitter



AGI-18

AGI CROSS-SCORE PROCESSOR



Processor Controls
- Control filter $h(t)$
- Frequency-shift value α
- Conjugation control (*)

AGI-19

AGI EXAMPLE SPECTRALLY SELF-COHERENT SIGNALS

Complex Modulation Format	Self-Coherence Frequencies	Conj. Self-Coherence Freq. (± 2 × Carrier)
BPSK	Symbol-Rate Mult.	Symbol-Rate Mult.
QPSK	Symbol-Rate Mult.	None
MSK, SQPSK	Symbol-Rate Mult.	± 1/2 Symbol-Rate ± Symbol-Rate Mult.
CPFSK	Symbol-Rate Mult.	Symbol Frequencies (h = multiple of 1/2)
FDM-FM	Pilot-Tone Mult.	None
DSB / VSB AM	None (complex repr.) 2 × carrier (real repr.)	0
SSB AM	None (stat. baseband)	None

AGI-20

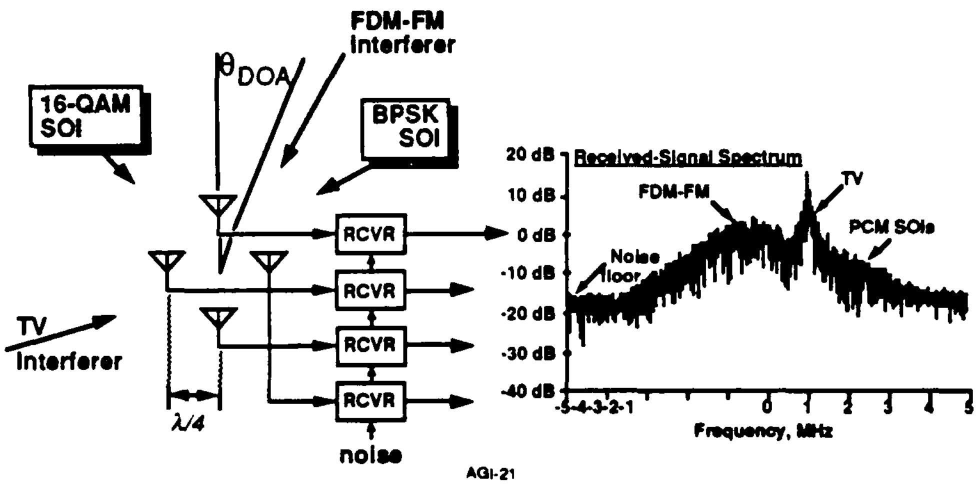
AGI

THE PROPERTY-RESTORAL APPROACH TO BLIND ADAPTIVE SIGNAL EXTRACTION

AGI

EXAMPLE RECEIVED ENVIRONMENT

Signal	Carrier	Symbol Rate	SWNR	DOA
16-QAM	0	3 Mb/s	15 dB	-45°
BPSK	0	4 Mb/s	20 dB	60°
FDM-FM	-500 kHz	-	30 dB	30°
TV	2 MHz	15.625 kHz	40 dB	-110°



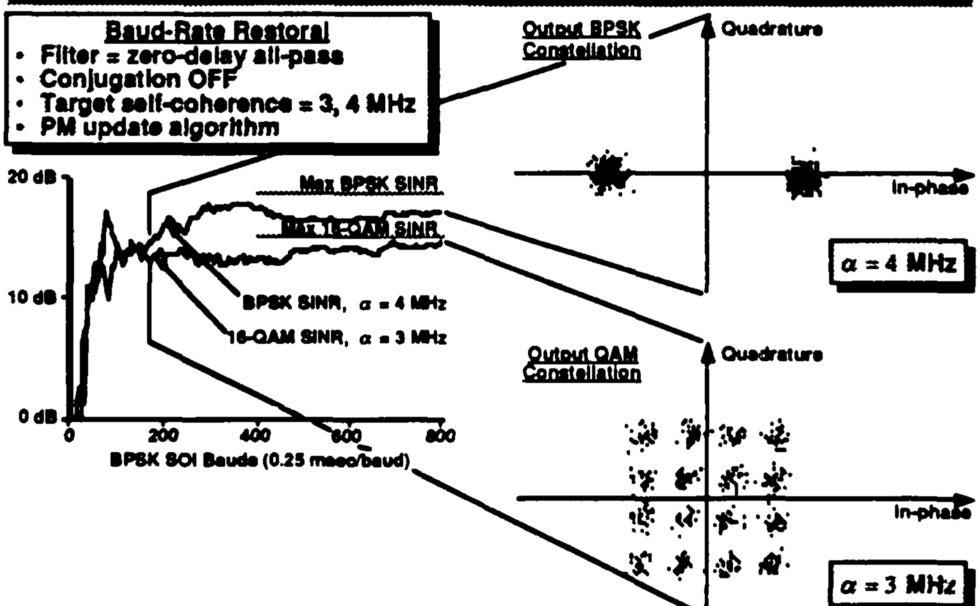
AGI-21

AGI

SELF-COHERENCE SORTING PERFORMANCE

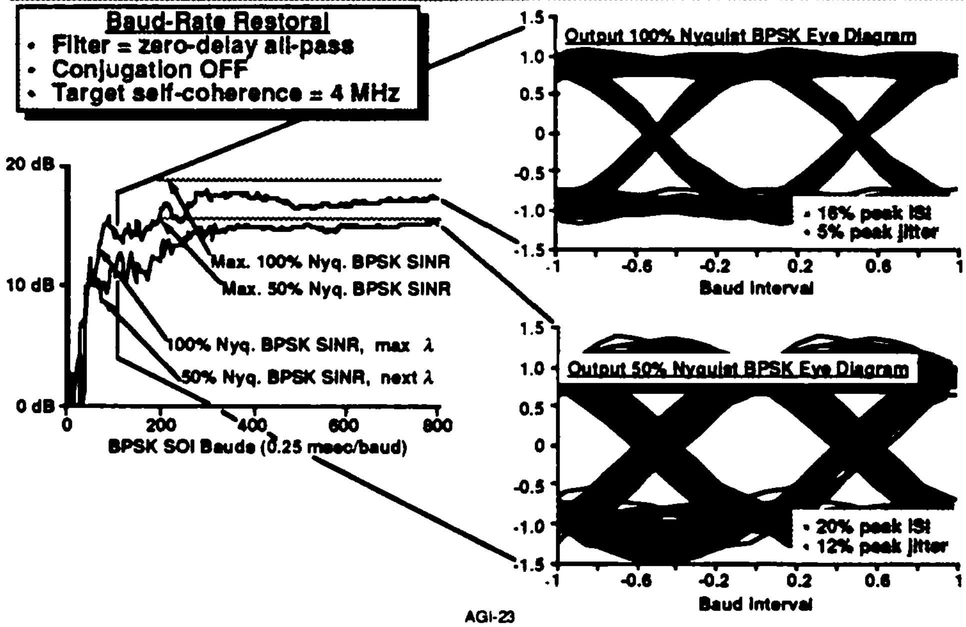
Baud-Rate Restoral

- Filter = zero-delay all-pass
- Conjugation OFF
- Target self-coherence = 3, 4 MHz
- PM update algorithm



AGI-22

AGI EIGENEQUATION SORTING PERFORMANCE



AGI-23

AGI PAM DEMODULATORS

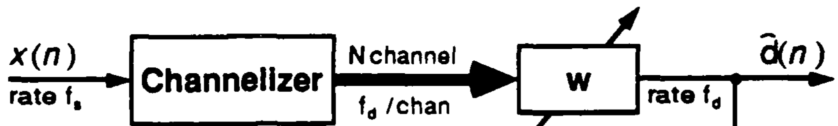
- Powerful extraction capability compared to LTI filter
  - Superior equalization of channel distortion
  - Superior nulling of narrowband SNOIs
  - Null-steering on wideband PAM SNOIs
- Very amenable to blind adaptation via modulus restoral
  - Insensitive to carrier offset
  - Delay ambiguity limited to unit-symbol
  - SOI properties much more exact
- Disadvantages: Susceptible to noise capture
  Requires knowledge of SOI symbol-rate

AGI-24

THE PROPERTY-RESTORAL APPROACH TO BLIND ADAPTIVE SIGNAL EXTRACTION

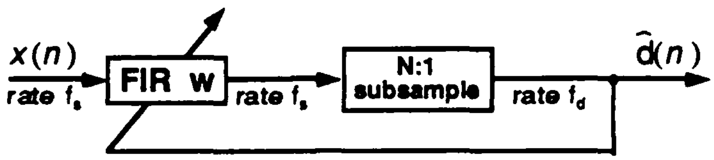
LINEAR-PAM DEMODULATOR STRUCTURE

- G eneral demodulator structure ($f_s = N \times f_d$)



- Time-domain (DMX)
- Freq-domain (FFT, TMX)

- **Equivalent structure for DMX / FFT channelization *(fractionally-spaced equalizer)***

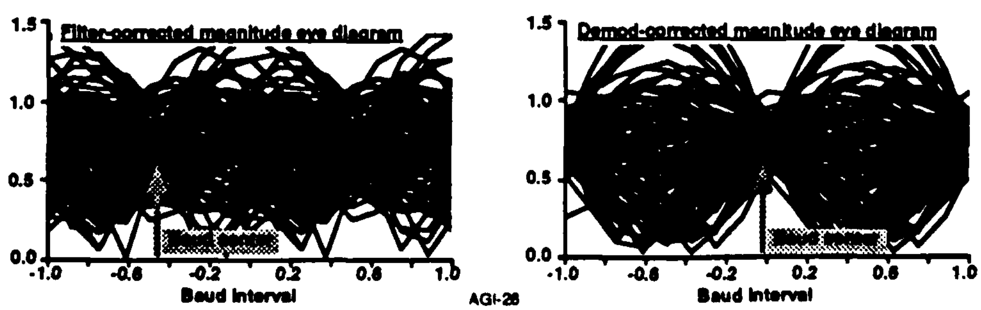


AGI-25

NARROWBAND DEMODULATION EXAMPLE

Environment
- 10% Nyquist QPSK
- 13 dB E/N
- 1/3 symbol timing error

Run Parameters
- 1000 symbol collect
- 2 samples/symbol (N = 2)
- 64-tap FIR filter/demodulator processor
- Static LSCMA adaptation algorithm



AGI-26

nal symbol sequence are generally much more "exact" – a PSK symbol sequence, for instance, *always* has a constant modulus (even if the PSK signal is frequency shifted), but the PSK waveform can have a high modulus variation if it is generated using a spectrally efficient pulse. The second advantage results from the superior performance of the optimal (nonblind) demodulator. The optimal demodulator reduces to the generalized matched filter in stationary interference environments; however, it has even more useful properties in *cyclostationary* interference environments. For instance, the linear PAM demodulator has the capability to *null-steer* on signals not of interest (or other SOIs) if they are also PAM with the same baud rate as the SOI.

This nulling property is illustrated in Slides AGI-28 and AGI-29, for an environment containing white noise; a particular kind of pseudonoise spread spectrum (PNSS) signal, which I refer to here as *PNSS modulation-on-pulse (PNSS-MOP)* or *phase-coded pulse-compression*, received at a 0 dB SWNR with a 16:1 spreading ratio; and a PNSS-MOP interferer with the same message rate and spreading ratio and a 30 dB SWNR. A PNSS-MOP signal can be interpreted as any PNSS signal where the code repeat rate is equal to the data rate of the pre-spread message sequence. In this case, the signal can be thought of as a linear-PAM signal, where the modulating sequence is the data or message sequence and the modulated *pulse* (rather than the message data) is spread by the code sequence. The message sequence can therefore be extracted directly from the received data using an FSE with an output data rate equal to the SOI message rate.

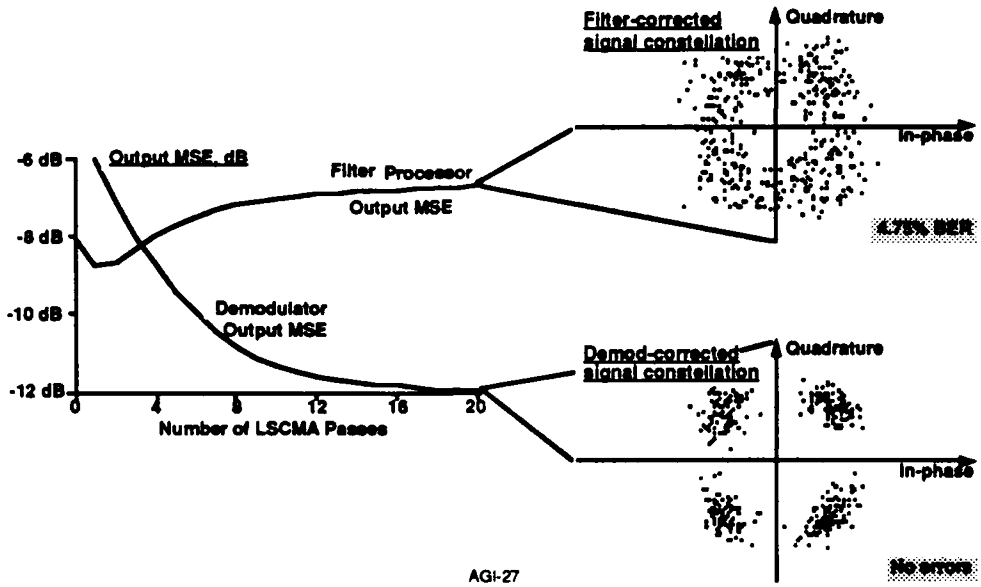
So, our received data consists of white noise, a PNSS signal at a 0 dB SWNR (about a 3 dB E_b/N_0 for the chip shaping assumed here) and a PNSS interferer at a 30 dB SWNR (33 dB E_b/N_0). This results in a signal-to-interference-and-noise-ratio (SINR) of about -30 dB, which is too low for the conventional matched-filter despreader to reliably demodulate the SOI – the SINR at the demodulator output would only be at about -17 dB in this case. However, as Slide AGI-29 shows, the FSE-based demodulator structure is able to increase the SINR of the demodulated message sequence to about +15 dB, easily high enough for the low-error detection of the message sequence. In fact, the demodulator has sacrificed one degree of freedom to null the interferer, and is using its remaining degrees of freedom to perform matched filter demodulation of the SOI. This is very similar to the null-steering and beamforming operations displayed by an antenna array when a signal and interferer arrive at the array from different angles of arrival.

Moreover, in this simulation the signal extraction was accomplished blindly, using an LSCMA. Thus it is also possible to accomplish this extraction without knowledge of the spreading sequence of the SOI.

To conclude, I'd like to discuss a promising area for future investigation of blind adaptive processing. In the area of processing structures, the most interesting areas to me are: development of general property-mapping approaches for exploitation of arbitrary SOI properties; generalization of the demodulator structures to allow extraction of basebands from more more exotic signal waveforms (such as general PNSS, nonlinear PAM); and extension of the multitarget algorithms to more exotic environments and processor structures. The appeal of the property-mapping approach is that is can be applied to any signal property, and that

THE PROPERTY-RESTORAL APPROACH TO BLIND ADAPTIVE SIGNAL EXTRACTION

DEMODULATION PERFORMANCE



AGI-27

WIDEBAND NULLING EXAMPLE

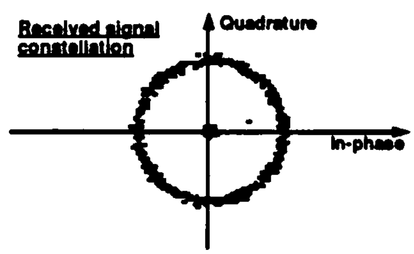
SOI Parameters

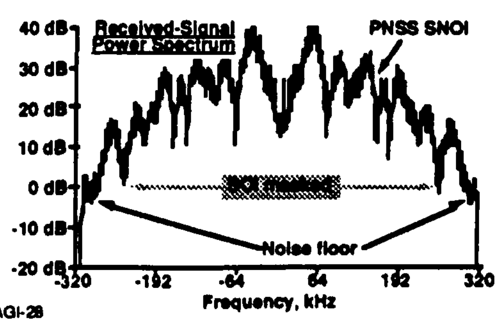
- 20 ksym/sec QPSK message seq.
- 320 ks/s QPSK spreading sequence
- 16-chip PN spreading code (MOP)
- 100% Nyquist chip-shaping
- 0 dB received SWNR
- No timing or carrier offset

SNOI Parameters

- 20 kb/s BPSK message sequence
- 320 kb/s BPSK spreading sequence
- 16-chip PN spreading code (MOP)
- 100% Nyquist chip-shaping
- 30 dB received SWNR (-30 dB SIR)
- 15-chip timing offset
- 32 kHz carrier offset

Matched filter fails here (MF output SINR ≈ -17 dB)

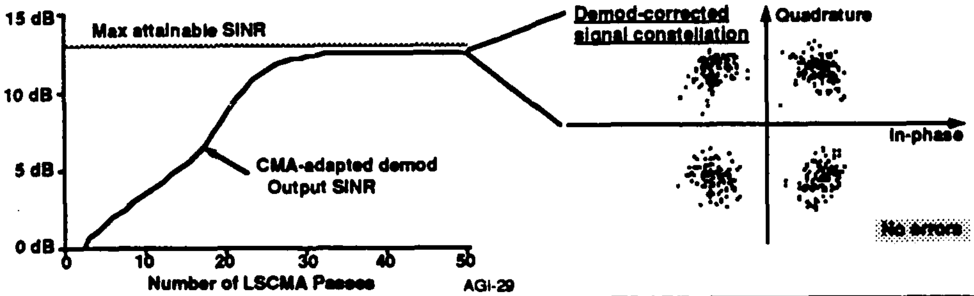




AGI-28

AGI EXTRACTION PERFORMANCE

Run Parameters
- 1 kbaud (16 kchip) collect
- 32 samples per baud (2 samples per chip)
- 64-tap FIR demodulator
- Static LSCMA adaptation algorithm



AGI-29

AGI PROMISING AREAS FOR FUTURE INVESTIGATION

- Processor structures:
  - Property-mapping structures
  - General demodulator structures
  - Multitarget demodulators

- Algorithm invention & development
  - General property-mapping algorithms
  - Combined POCS/PM and blind property-mapping algs.
  - Genetic algorithms

- Theoretical development
  - General convergence analysis
  - Extension of ML estimator

AGI-30

it immediately results in a monotonically-convergent adaptation algorithm. In particular, I believe that some very powerful new algorithms and theory are going to result by linking this approach with the methods of Youla [12] and Cadzow [14], and I strongly urge anyone interested in this area to familiarize themselves with this work. I also believe that it should be possible to use the nulling property of adaptive demodulators to develop a multitarget demodulator that can separate and sort signals in single-sensor systems. Between all of these developments, I think the time is growing near when we can actually address the co-channel sorting problem posed by Dr. Pickholtz in yesterday's session.

References

1. J. R. Treichler, B. G. Agee, "A New Approach to Multipath Correction of Constant Modulus Signals," *IEEE Trans. ASSP*, vol. ASSP-31, pp. 459-472, April 1983.

2. J. R. Treichler, M. L. Larimore, "New Processing Techniques Based on the Constant Modulus Algorithm," *IEEE Trans. ASSP*, pp. 420-431, April 1985.

3. B. G. Agee, "The Least-Squares CMA: A New Technique for Rapid Correction of Constant Modulus Signals," in *Proc. 1986 Intl. Conf. on ASSP*, 1986.

4. R. P. Gooch, J. Lundel, "The CM Array: An Adaptive Beamformer for Constant Modulus Signals," in *Proc. 1986 Intl. Conf. on ASSP*, 1986.

5. B. G. Agee, "The Baseband Modulus Restoral Approach to Blind Adaptive Signal Demodulation," 1988 Digital Signal Processing Workshop, September 1988, Tahoe, CA.

6. B. G. Agee, "The Property Restoral Approach to Blind Adaptive Signal Extraction," Ph.D. Dissertation, University of California, Davis, CA, 1989.

7. B. G. Agee, "Blind Separation and Capture of Communication Signals Using a Multitarget Constant Modulus Beamformer," in *Proc. 1989 IEEE Military Communications Conference*, 1989.

8. A. Benveniste, M. Goursat, "Robust Identification of a Nonminimum Phase System: Blind Adjustment of a Linear Equalizer," *IEEE Trans. Automat. Contr.*, vol. AC-25, pp. 385-399, June 1980.

9. D. N. Godard, "Self-Recovering Equalization and Carrier Tracking in Two-Dimensional Data Communication Systems," *IEEE Trans. Comm.*, vol. COM-28, pp. 1867-1875, November 1980.

10. J. Lundel, B. Widrow, "Application of the Constant Modulus Adaptive Beamformer to Constant and Non-Constant Modulus Algorithms," in *Proc. Twenty-First Asilomar Conf. on Signals, Systems and Computers*, 1987.

11. B. G. Agee, "Convergent Behavior of Modulus-Restoring Adaptive Arrays in Gaussian Interference Environments," in *Proc. Twenty-Second Asilomar Conference on Signals, Systems and Computers*, 1988.

12. D. C. Youla, H. Webb, "Image Restoration by the Method of Convex Projections: Part 1 – Theory," *IEEE Trans. Med. Imaging*, vol. MI-1, pp. 81-94, Oct. 1982.

13. M. I. Sezan, H. Stark, "Image Restoration by the Method of Convex Projections: Part 2 – Applications and Numerical Results," *IEEE Trans. Med. Imaging*, vol. MI-1, pp. 95-101, Oct. 1982.

14. J. A. Cadzow, "Signal Enhancement – a Composite Property Mapping Algorithm," *IEEE Trans. ASSP*, January 1988.

15. B. G. Agee, S. V. Schell, W. A. Gardner, "Self-Coherence Restoral: A New Approach to Blind Adaptive Signal Extraction Using Antenna Arrays," *IEEE Proceedings*, May 1990.

16. W. A. Sethares, G. A. Rey, C. R. Johnson, "Approaches to Blind Equalization of Signals With Multiple Modulus," in *Proc. 1989 Intl. Conf. on ASSP*, pp. 972-975, 1989.

HALL: Thank you very much. We'll have a ten minute break and get back for the Question and Answer Session. [PAUSE]

RAYMOND PICKHOLTZ: I actually have two questions. First, one for John Cioffi on the very excellent talk he gave about combining equalization, pulse shaping and coding to try to get closer to the predicted capacity of the kind of channels we're talking about. I spoke to him briefly during the break, and one of the questions I have is, "What is the relationship between this and some other work that's been done for a very special channel, namely the magnetic recording channel which is, of course, a very special channel because it essentially has only two levels that you deal with?" In particular, Jack Wolf at UCSD and Chris Heegard at Cornell have done some constructive coding – which I don't think, John, that you specifically talked about in your talk – some constructive coding for that magnetic recording channel to combine both coding for distance properties and run length limited coding for the purposes of fixing up the intersymbol interference problem. It was my understanding that they had made some significant advances, especially with respect to getting closer to the capacity of that magnetic recording channel.

CIOFFI: I, too, have worked in the recording industry; in fact, when I was at IBM that was my job function. If I look at my current base of funding, actually, it's a majority in that area from both IBM and the state of California. So I have looked at this problem for quite a while; I'm intimately familiar with the work by Wolf and Heegard as well. When you come back to this particular slide, the same type of situation arises in what they're doing in that you have some kind of coding gain for the system, and again they are looking at the combined equalization and coding problem. You correctly surmised, Ray, except they have as well, or we have in magnetic recording a binary constraint at this point in the network. [FOIL #4] It's two levels only. The actual clock rate can be about as high as you'd like. The practical constraints on the write clock in recording are not really limiting in this particular situation. But it's only two levels because of the hysteresis effect which occurs in magnetic media. Now when you're left with the two level constraint like that, both of the techniques that I presented here will fail in that particular area because of the fact that they require more than two levels on the input to the channel. Both the Tomlinson precoder and multitone approaches are basically pretty widely dispersed over the entire transmit range of symbols that you could have.

Now for that particular channel the thing you have to look at is this equalization gain.

Given you have a binary constraint, the only way to really get your density up is to increase the clock rate. It has to increase – if you're going to use a code it has to go up above your actual data rate or data density for the disk. When you start running up symbol rates in systems you have to take a very significant hit in equalization loss. You can try to model the channel by an increasingly more complicated series of partial response polynomials at higher and higher densities, and try to design the codes for those situations.

I have a fundamental point of difference with both the Heegard and Wolf works. They're not really here and they would argue with me at length on this, as they have in the past, but when you start running up the clock rate on the system there is a very significant hit in equalization loss that you have to pay before you can add on a coding gain. I have yet to see a code that anyone has come up with where we cannot design an uncoded system that would beat it significantly. So it's still an open problem, as far as I'm concerned.

I have one student in my group who works – he's well aware of this difficulty – and he came up with a code that did as well as an uncoded system, which is really nothing to brag about. But in this sense where you have to increase the clock rate, is the only way to code, you pay this equalization loss penalty. Most of the coding approaches to that have, one way or another, avoided including that penalty in the calculation. So I say it's an open problem at this point. I don't pretend to have it solved; I don't think anyone else has it solved at this point. The work that seems to be going the best along this line is by a couple of guys, Paul Siegal and Raznik Karabed at IBM Research in San Jose, who understand this equalization loss issue. Of course, the gains that they predict for their codes are far less than what you see projected on these other codes that you refer to because they are considering that particular loss.

My gut feeling is that the actual way to do it is to basically start to border on the range of an FM signal in magnetic recording. In effect, you're doing FM with the square wave. To use the thing that you'll incur is that in most FM transmission paths you really don't have such a severe ISI induced in a transmission path. So my gut feeling is that's the way to go, but I can't say that it really is ....

PICKHOLTZ: I agree with what you're saying, John. The question I had is that: since, in the magnetic recording problem, you do have this severe constraint of two-level signals, and therefore the only way you can go is by some run limited coding pattern, it would seem that your problem – you know, the general problem here – would be easier, but that you might be able to take advantage in the shaping characteristics of using some kind of run length limited characteristics in a multilevel signal, rather than in a binary signal, as you have in magnetic recording.

CIOFFI: That's an interesting – a number of people have raised that issue. I haven't seen any work along those lines yet, but I would agree that if you could solve the binary problem then it might well apply to a multilevel situation where you might have, for clocking purposes or some other reason, a constraint on the number of levels in your transmit – say 4 or 8 levels – rather than a continuum of levels which span the entire interval. I still see that as, it's in the same framework as you guessed, as the other problem, but I don't believe it's been solved just yet. But I think if you can solve one, you can solve the other problem. It's a very good point that you raised, and that problem is

just intrinsically harder at this point to solve because it has an additional constraint, with respect to the one that I was considering.

HALL: Ah, yeah, Dr. Peile ....

PEILE: The question is for John Cioffi again, and he knows the question .... It would help to have the slide up with the multichannel model up so people could see it. While John is finding the slide I'll just say that he mentioned the need for the possibility for interblock coding, rather than have a bank of independent decoders, one for each channel. I had a look at that problem about a year ago, I think about a year ago – I'll wait a second until the slides are up. [PAUSE] That was the one – that'll do at any rate.

CIOFFI: My comment where I put up this summary ....

PEILE: Yeah, the comment about interblock ....

CIOFFI: ... about interblock coding ....

PEILE: Yeah, but that's the diagram I wanted to look at ....

CIOFFI: OK, sure.

PEILE: I looked at the possibility of having one code across all the channels rather than having several independent decoders, presumably identical, I don't know. But I had one code across sets, and instead of making the signal-to-noise ratio constant into the decoders I have tried to arrange the power levels so the signal + noise into them was constant, more or less. I then took the code and adapted it to those conditions, so that the output error in the decoder was about ....

CIOFFI: The output error rate was what?

PEILE: Well, the point is that as long as the code was adapted to the input conditions they didn't have to be constant on each channel. By doing that I found at least some sets of conditions which seem to win; but it's not clear to me when that is a winner and when you really want to have a multichannel in the signal-to-noise ratio into independent decoders to be constant. Obviously the drawback with that is throughput delay, having to unset the codes on each channel.

CIOFFI: Right, well, the reason for going to the crossblock coding if you have n tones here, you go across the block with the code rather than put a trellis code on each one is delay. That's the reason for citing that. You say you've looked at the problem and have looked at a block code – if I remember correctly from your paper ....

PEILE: It's a finite number of channels where I had a ....

CIOFFI: ... over a finite number of channels ....

PEILE: ... finite number of bits ....

CIOFFI: Yeah, and then you use a slightly different constraint. It may be possible to have other techniques that basically – I'll also get to this – are as close to capacity as you're going to get within the multitone framework. I wouldn't dispute that you can use slightly different design criterion and come up with the same type of result here.

What I was referring to with the crossblock coding is when you use a block code, there are basically some edge effects that occur that you either have to get rid of using zero stuffing or overlap save methods or some equivalent to that. What I think I was referring to is kind of designing almost a convolutional or trellis code that goes down this set and into the next block, and into the next block after that ....

PEILE: I believe that's possible ....

CIOFFI: ... and you have different numbers of levels on each of these tones. This is a problem that someone actually raised at Telebit – asked me if I knew a solution to this. They were very interested in designing a code

along those lines. So that's why I put up the interblock coding. So I was more referring to a convolutional type thing, but would agree that your method probably would get to this same performance level.

PEILE: Well, I think you could almost have a reverse concatenation, where you have the inner code as a block code across the channels, and then a convolutional trellis going along in time across the channels.

CIOFFI: Right ...

PEILE: ... Which is the other way around the normal ....

CIOFFI: ... and the objective again is the delay, get the delay down, because into multitone schemes are notorious. I guess I have the "T" and the "L" transposed here – I got "mutlichannel" ....

PEILE: I'd also put out that you used to call them "eigenchannels," which I thought was a great name and now ....

CIOFFI: Right, well, I was just trying to simplify. In fact, I put vector here, I didn't really put tones and, of course, the best thing to do is not put a DFT-type thing there. There are better approaches – that's what I call "vector coding," which will give you, for a given fixed blocklength, a higher gain than the DFT approaches will. But you'll have to know what the channel is in designing the vectors, whereas the DFT-type vectors are independent of the channel. They converge to the same thing as the blocklength goes to infinity. But your – I wouldn't dispute your point at all, that there is, you know, probably other ways of designing the details of the code so that you would achieve the best performance. And what I meant by interblock coding was to get the delay down, but go from block to block in a finite-state type sense as opposed to ....

PEILE: OK, that was ....

CIOFFI: ... fixed block approach.

PEILE: I think Vedat disagrees with both of us, but we'll find out this afternoon ....

CIOFFI: Pardon?

PEILE: I think Vedat disagrees with both of us, but we'll find out this afternoon.

CIOFFI: Yeah, well, we've had some agreements and disagreements over the years, but I welcome that .... Are you going to ...?

BILL GARDNER: I think there might be a way to reinterpret your approach, and if I'm right it may show a link to some earlier work. I think you can view, say, the convolutional encoder and pulse shaping filter together as a linear periodically time varying processor. Also you can view the received filter, linear demodulator and linear decoder as a periodically time varying processor.

CIOFFI: OK.

GARDNER: And Graef – I'm not sure of the pronunciation, it's G-R-A-E-F – in 1983 NATO-ASI (Advanced Study Institute) did a paper on "Joint Optimization of Periodically Time Variant Transmit and Receive Processors," that I think may in fact be related to what you're looking at, by separately designing the encoder and a pulse shaping filter of the transmitter.

CIOFFI: I'm not generally familiar with that work, but we did try that approach of jointly designing both together, and we're not successful that way. This way you'll get as good a code as you're going to get. Use either of the two techniques that I presented. In fact, one of the agreements with Vedat was last night at dinner. We agreed that that approach is probably the wrong direction, that you cannot get the gains that you'd like by concatenating the two together. To the best of my knowledge, no one has solved that problem. I'm not familiar with this particular work, but based on several intelligent grad

students beating their heads at it, as well as other people I know who've looked at that particular approach, have not been successful. I just have a gut feeling that maybe that's not the right direction to go. But given that I haven't seen this '83 NATO study, I can't say for sure ....

GARDNER: OK, and I certainly can't either, but ....

CIOFFI: ... just exactly what the gains were.

GARDNER: I want to look at it, anyway ....

CIOFFI: Well, certainly it would be a welcome input to this study. One more question in the back?

BART RICE: I had a question for both John and Brian. I was interested in your precoding polynomials. If I understood them right, there were two that *seemed conspicuous* by their absence, and those were the two that are associated with the duo binary and modified duo binary. Duo binary should be $1 + D$ and modified duo binary should be $1 - D^2$. Neither one of them was there, and I wondered if there was some reason for that. Also, how you got the polynomials you got.

CIOFFI: OK, the reason the $1 + D$ and $1 - D^2$, which are the dual binary are not shown is because they're exactly the same as the $1 - D$ channel, the AMI channel. The $1 - D^2$ is basically 2-interleaved, $1 - D$'s on the evens and odds, so you'll get exactly the same result. Because $(1 + D)(1 - D)$ is just $1 - D^2$, and it's going to be the same result. In the $1 + D$, the plus or minus sign really has no effect, it just shifts the notch from Nyquist frequency down to DC, and they're completely symmetric. So the results would be exactly the same as the $1 - D$ channel.

Now where I got – the second part of your question – where I got these partial response polynomials from is my background in storage. They are a set of polynomials called "extended partial response class," that people like to use to approximate a magnetic storage channel at increasingly higher and higher densities. The way you get higher densities is to take the $1 + D$ factor and increase the exponent on that. So that's just one particular class, but what it represents is a series of increasingly low pass channels with more and more severe ISI, which I used as a representative example to illustrate the point about the distance to capacity increasing when you have this more severe ISI in the channel.

RICE: And the question for Brian was: "On your multilevel, your multi-amplitude, instead of the constant modulus, your multimodulus algorithm, was that used to try to equalize to the signal or to try to recognize the modulation type?" I was thinking about this after we talked yesterday. I think that the approach we took was to try to equalize just using the constant modulus algorithm, and then try to recognize the modulation type using the amplitude distribution. So are you trying to equalize also using the multiamplitude distribution?

AGEE: Yes. More exactly, I'm trying to extract that baseband sequence from the noise and interference by using its multiple modulus. The reason why I asked you that question yesterday was, one of the unanswered questions is, "What can you do if you don't know what those modulus levels are?" I'm curious to see if there's some way that the multiple-modulus approach can be combined with the clustering approach that you're using, to look at the more general case where we just know it's a QAM signal but we don't know what its levels are at. But, yes, it was basically an equalization approach.

RICE: I think that the ... we thought

about that and abandoned it because for two reasons. I think (1) we thought that if there was any performance margin over the ordinary constant modulus algorithm that it would be marginal; and (2) it depended pretty carefully on where you placed the rings, the amplitude rings, because – especially if you've got things coming from a demodulated *IQ* point that came in between two rings – it seemed like the probability of error was too great of sending the error signal in the wrong direction. So we abandoned that fairly quickly, but maybe too quickly.

AGEE: Yes, I think there is a potential for a problem there, and that's again one of the reasons why I'd like to see the technique extended to do some kind of adaptive determination of those modulus levels. However, with respect to your comment that you think that it's going to have marginal performance improvement over the CMA, I disagree with that because one of the notable properties of the constant modulus or the dispersion directed algorithms is that they converge very slowly for QAM signals. My interpretation of why that's happening is that there is not a close match between the property that we're trying to restore and the signal that we're trying to extract. We're trying to use a constant modulus property, but the signal doesn't have a constant modulus. So we have a residual error in our modulus variation, which slows adaptation. The idea here is that the multiple modulus approach is going to remove that error and it should accelerate our convergence. The other thing is that I developed this approach in tandem with the least squares algorithm, which should further accelerate convergence of the algorithm. I have to say that most of my results in this area have been theoretical. I have been able to prove monotonic convergence of the algorithm; I have also been able to prove, for certain kinds of environments, that the stationary points of the algorithm are going to correspond to signal capture. However, the multiple modulus algorithm has not at this point been simulated, so I don't have any hard experimental results. But the theoretical results look very good. My experience looking at the rapidly converging algorithms is that a lot of problems that are associated with the CMA just go away when you start looking at the rapidly converging algorithms – problems with misadjustment and things like this.

HALL: If you formulated that in a least squares, couldn't you go ahead and add the exact ring positions or the exact modulus values to the least squares formulation, and just adapt on those also?

AGEE: I'm not quite sure I understand your question. I probably shouldn't have used the term "least squares." It's the terminology that I developed ....

HALL: Oh ....

AGEE: It's not really a least squares algorithm, it's more of an alternating projection algorithm, and one of the projections is least squares and the other one is a projection onto the desired signal set. Given that, could you rephrase your question, or did that ...?

HALL: Mine pertained to a least squares formulation. You could just leave the modulus, you know, like for QAM or 16 QAM you have three modulus positions, right? You could leave those as unknowns and solve for those in your least squares formulation.

AGEE: Ah, possibly. You have to be careful though to avoid a trivial solution. The algorithm might just set all the modulus levels equal to zero, and then set the processor weight equal to zero, and then the cost function becomes equal to zero. So you have to be a little careful.

I encountered the same problem with one of the other algorithms that I call the "adaptive modulus algorithm," which works on this almost periodic property of the modulus. I was able to overcome that by dividing through by the mean square error of the modulus itself. Unfortunately that's a lot more difficult to do for the multiple modulus algorithm. If I kind of wrote it out for you it would be easy to see, but it's just that the minimization problem is very difficult if you turn it into a ratio of two time averages instead of just a single time average. And then, John ....

TREICHLER: Was he just trying to get out of the rest of the answer, or did he just hand me the microphone? I wanted to respond a little bit to Bart's comment. I mentioned to Brian a few minutes ago that Monty Frost was the one who came up with that known modulus scheme, exploiting the case where there's a known deterministic modulus variation. The varying envelope has no information, but it destroys the constant modulus property. Also, in a little internal tech note he wrote in '82, '81, '83, sometime back in there, he suggested this multiple modulus scheme for working what were in the trade known as V.29 modems, which at that time was considered to be a very hard problem. We actually did a little simulation work on it, not a great deal, and found that in fact it worked very nicely and it did in fact work faster than pure Godard or CMA. The reason we abandoned it was not that it didn't work, but that it didn't converge order of magnitudes faster than Godard. On the other end of the scale it was a little bit more sensitive in terms of where you chose the decision levels. But the big killer is we wanted to generalize upward. We wanted to go not just at 16 QAM, but 32 and 64 and 128, and it gets messier and messier as you try to pick these decision circles closer and closer together. It becomes more and more sensitive and we said heck, so we ended up with sort of two ends of the scale – Godard at one end that was excruciatingly simple, and then decision directed at the other end to handle the very high constellations once Godard had stabilized.

AGEE: It just occurred to me that one possibility is to do some kind of a variation on the reduced constellation algorithm. We might still be able to get some performance improvement if you choose a set of rings which is fairly close to certain concentration of rings. You know, deliberately choose a smaller number of rings. I guess I didn't quite answer your question. My answer is I don't know, I think it's an interesting idea to try to adapt the modulus, and again that's why I was asking Bart Rice yesterday if he looked at any kind of clustering algorithms. I'd like to see something on that.

SCHOLTZ: Thank you, Dennis. Since I don't work in this area I'll probably have to ask some dumb questions for the rest of the crowd that's being quiet. I don't know they're staying quiet. [LAUGHTER]

You're talking about, for example, the constant modulus algorithm, and you have this function that you'd like to minimize – I've forgotten exactly what is the (modulus - 1) quantity squared, or something like that. I can think of all sorts of other functions which you would like to use. Now, why have you chosen that particular one? Is that a complexity issue or is it a performance issue, or has anyone looked at all the other possible functions that you could use to adapt on? [PAUSE] Is that a dumb question?

AGEE: No, it's not a dumb question; in fact, a large part of my research has been to try to determine which of these is best. There are several constant modulus cost functions;

at ARGOSystems, we generated what we call the "general $p-q$ constant modulus cost function." One of the questions that has remained unanswered is which one is best. And there are different ways you can look at "best." From an implementation point of view I believe the 2-1 and 1-1 constant modulus algorithms are easiest to implement, and I think John is going to probably want to say something on that where the $p-q$ cost function is given by $< (|y|^p - 1)^q >$.

From a performance point of view or from an implementation point of view the so-called $p-1$ algorithms, which take the *magnitude* of the signal to the p^{th} degree -1, tend to be the easiest to implement. However, the 2-2 CMA is the one that's been analyzed the most because it involves just second and fourth order moments, so people can say a lot about it. The 1-2 algorithm has still other advantages – I personally like that one best because it has a very interesting form. You can formulate it so that it looks a lot like a demod-remod type of algorithm. Also it fits very well into the property mapping viewpoint, which allows you to develop rapidly converging adaptation algorithms. Also, you can formulate the blind signal estimation problem in such a way that you can derive a maximum likelihood estimator of the signal, if you have an antenna array and you've got temporally wide interference, and you can show that the optimal algorithm optimials the 1-2 cost function. So in some cases that algorithm appears to be best. On the other hand the analysis work that I've done shows that all of the $p-q$ algorithms have stationary points that correspond to capture of the signal as the collect time grows to infinity in certain environments. That's kind of a really interesting result. If you do a little bit more analysis, it looks like the 2-2 algorithm is possibly a little bit more robust if you don't know exactly what that signal feature is. So I guess my answer is is that there are a lot of different criteria, and depending on which criteria you use one algorithm is better than the other.

TREICHLER: Let me amplify with one minor comment. We've formulated that general $p-q$ description because in fact there was a great deal of argument in the beginning. For the 2-2 case, even I could differentiate the cost function, and so that's one of the reasons why we use that. My interest in the 1-1 algorithm is you're taking a magnitude, subtracting, and taking another magnitude which, if you come in with 8 or 10 bit data, means that from a numerical point of view you're not doubling the wordlength or quadrupling the wordlength. So that was my interest originally in pursuing the 1-1. We sort of ended up with these ones where you square once and magnitude once as a compromise between understanding the 1-1 and 2-2 algorithms and making them reasonable to implement. You can do other things like magnitude of the log of the magnitude, and just all sorts of – I mean, you can do almost anything. Nobody really knows anything to say what's better than the other, other than the ones Brian has looked at in detail.

AGEE: I had one last comment and that is that if you formulate the blind adaptive problem as a property mapping algorithm, then you basically get one cost function form. So a lot of that problem goes away. Of course, there might still be a better form out there for a given implementation, so you shouldn't necessarily depend on it. But a lot of that does go away. So I'm kind of settling on the 1-2 algorithms for that reason.

SCHOLTZ: I have one other comment. I've just read a few papers on these subjects, and I always thought that it was very inter-

esting in Widrow's algorithms, for example, that you need the signal that you're trying to get before you can do anything, which leads you to the decision-directed ideas. The beauty of the CMA algorithm is that you really don't need to know much about the signal in order to be able to build this device.

I'd like to hear a little bit more about some of the other reference signals and how you would apply them. How would you use cyclostationary-stationary, for example? I think you mentioned that in one of these blind equalization schemes. Is that a leading question? [LAUGHTER] In other words, how much information do you really need to know to apply that particular approach?

AGEE: I guess I'd say I should probably talk about that off-line, because I could go on to that for awhile, and I wouldn't want to do that!

SCHOLTZ: Could you answer just one specific question. How much do I really need to know about the modulation format in order to make that system work? In CMA all I need to know is that it's a constant envelope. How much more do I need to know to handle the cyclostationarity approach?

AGEE: I actually think I can do that in a few viewgraphs. Essentially all you need to know is one or two properties of the signal, such as its baud rate or the fact that it's a BPSK signal or a QPSK signal. Let me just put up a viewgraph to demonstrate that ... it'll take me a second to find it. You might want to ask someone else a question while I'm trying to find them, just to keep some continuity.

TREICHLER: There's no doubt somebody in the room wants to know why when John Proakis put up his description of all the blind algorithms that existed, not mentioned on that list was constant modulus, yet half the rest of the crowd in here seems to refer to it. It's exactly the same algorithm as Godard's. The two names came from two different application areas. John Cioffi and I were talking about this the other day, and there were at least four or five people at four or five organizations at the same time in the late '70s kicking around the same problem, most of them unknown to each other. In the modem world, as they went from QPSK and 8-PSK on up into QAM, there was a real question of how in multi-user networks to attain equalization of an equalizer without having to go back and ask for training again from the transmitter. People at AT&T Bell Labs were working on it, people at IBM were working on it. The group I was in wasn't working on that; we didn't even know about those people. We were trying to solve another problem for a classified customer, and that was trying to recover FDM/FM signals from a world of noise interference and multipath. In our case it was a really simple - well, it took us years to figure it out, but once you know the answer, everything is simple. We went in looking at all sorts of maximum likelihood, channel modeling, channel estimation, and all sorts of things like that. Finally, literally late one night, I said, "Hmm, we know exactly only one thing about this signal, and that is that it is FM, and it ought to have a constant envelope. I wonder if you can use that?" OK, this was in the late '70s sometime. That's exactly the same time the work was going on at Bell Labs and IBM and other places. It got published on that side and known as Godard's algorithm, because he's the first one who got it into a conference and then into a paper. In our community it was classified for a couple of years and finally bubbled up in the early 1980s, and it all depends on which community you came from as to what you call it.

AT&T is now looking at it in the microwave modem area as well. But they are the same thing.

PROAKIS: May I just add to that, John? Seymour Stein and I and a couple of his people at Stein Associates in the early '70's were doing blind equalization for DPSK signals, long before the term "blind equalizer" was coined. So I have a feeling that the work in the classified literature probably long predates some of the things that were subsequently published by Sato, Godard and others.

TREICHLER: I actually believe that Gauss did it as part of a [LAUGHTER] classified document.

SCHOLTZ: Maybe it's time for some of you older people in the field to sit down and write a history of equalization or something *like* that.

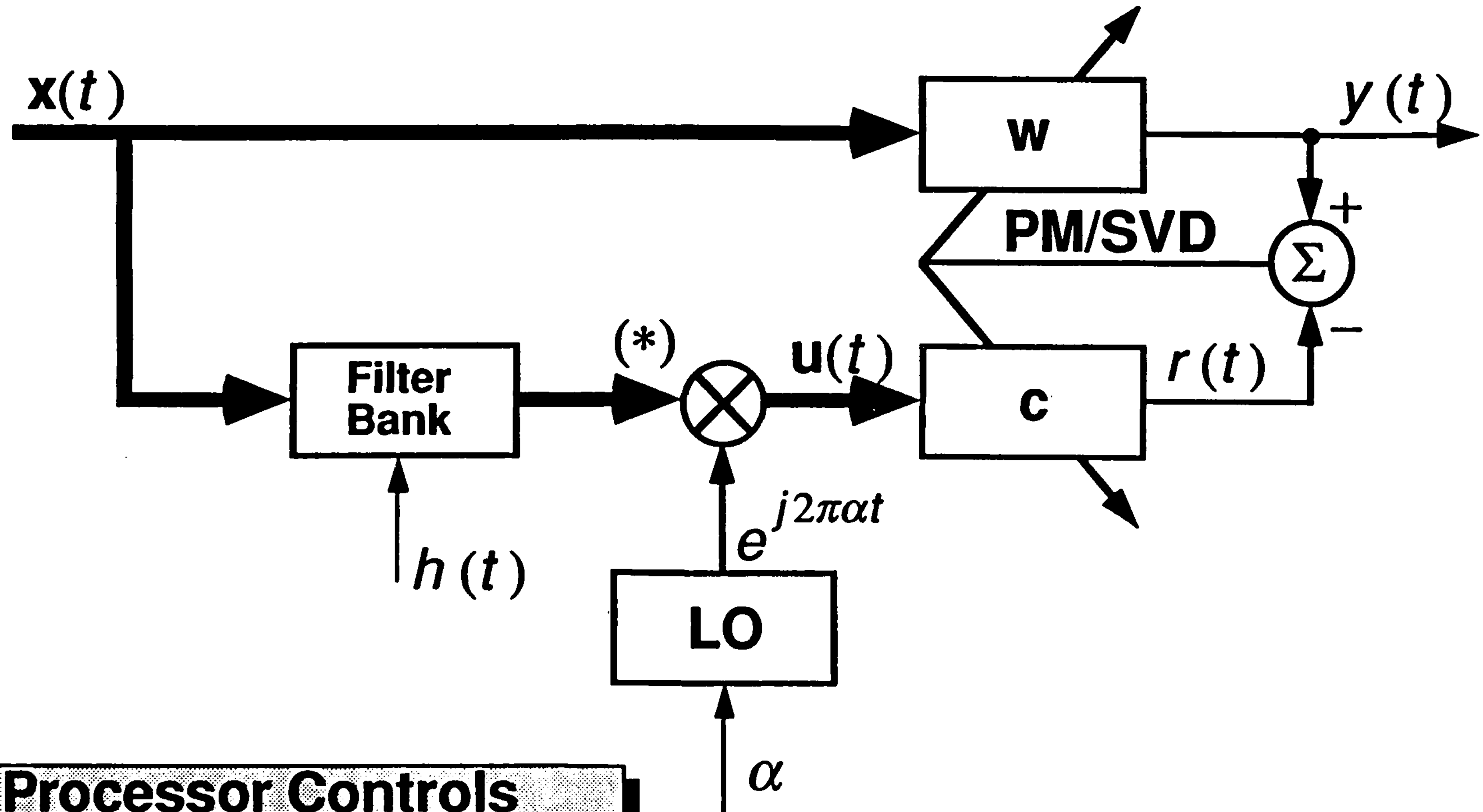
AGEE: OK, Slide AGI-19 shows the block diagram of the cross-SCORE processor, which is one of the self-coherence restoral processors that we've been investigating at UC-Davis. The processor path takes the signal, and passes it through a beamformer to get our output signal. The reference path forms a crude reference of the signal of interest. The main controls consist of a filter bank, which can be a delay line, or some kind of a filter, or a straight line, for instance for Nyquist-shaped signals. The filter output signal is then multiplied by a sinusoid and optionally conjugated, and then passed through what we call the control beamformer. The only two parameters that really determine the cyclostationarity property that you're looking at is α, the frequency shift parameter, and this optional conjugation. All that you really need to do is look at the various signals that you're searching for and figure out where are they self-coherent, and what kind of cyclostationarity they do have. Almost any PCM signal, for instance, is going to have cyclostationarity at its symbol rate, or it's going to have self-coherence and multiples of its symbol rate. So in order to implement this algorithm, all that we need to know is what the symbol rate of the signal is, which we have in a lot of applications. The algorithms that we've developed right now can do a pretty good job of detecting the signal if we know the baud rate of the signal of interest to within 1% accuracy. The additional parameter needed to configure the processor is the conjugation control, which is dependent on whether or not we want to exploit what we call the "conjugate self-coherence feature" or the self-coherence feature of the signal.

The optimal processor essentially solves a generalized eigenequation. The rank of that eigenequation tells you the number of signals that are self-coherent at the target α, while the eigenvectors in the signal subspace of that eigenequation correspond to capture of each of the signals in that environment under pretty general situations. In Slide AGI-22, we set $\alpha = 4$ MHz and basically captured a BPSK signal and ignored everything else; setting α to 3 MHz caused the algorithm to capture the QAM signal and ignored everything else. In Slide AGI-23, we replaced that 16 QAM signal by a BPSK signal with a different roll-off, but the same baud-rate at the first BPSK signal. As this slide shows, one of the modes of our eigenequation captured one of the signals; the other mode of our eigenequation captured the other BPSK signal. So we don't need to know much about it.

HALL: Let's see, I had a question for John Treichler. Something I asked him months ago was on adaptive IIR filters. Did you bring anything on those at all or ...? We only have

AGI

CROSS-SCORE PROCESSOR



Processor Controls

- Control filter $h(t)$
- Frequency-shift value α
- Conjugation control $(*)$

AGI-19

20 minutes left; I don't know if you can go into it.

TREICHLER: I have some viewgraphs, I can do it in 5!

HALL: OK, that'd be great. I'd really like to see that.

TREICHLER: Nobody believes that! I can hear the snickering from the crowd. [LAUGHTER] Watch the watch!

It's really a pretty easy topic to talk about because it hasn't gotten very far. I'm going to repeat Brian Agee and make a complete fool of myself fumbling through viewgraphs here. [PAUSE] [I found the one to start with, I've dumped three on the floor and] ... [VIEWGRAPH #11].

OK, I've organized this sort of the same way. What is IIR adaptive filtering? Who cares? What was the promise? Why do people want to look at it?

The idea of IIR filtering and the reason why a bunch of people have worked on it was that here was another methodology that promised to considerably reduce the computation in equalizers or adaptive filters when you were trying to model resonances or compensate for spectral nulls. If somebody has built a signal and transmitted it through a channel that had spectral nulling, for whatever reason, if I could do just the right IIR filter I could compensate for that. If in fact I was doing, say, an adaptive line enhancer, as they call it in the sonar world, where I'm trying to look for resonances or very narrowband signals and I would like to model those as little resonances, then maybe I'd like a filter that is built that way.

The other thought is that perhaps I can get filters with very, very long impulse response, but with very small dimensionality in terms of the number of coefficients. It's the general rule that when one adapts fewer coefficients rather than more, then everything gets better. So there was a fair amount of work done on this. I haven't even begun to touch on all the people who did various things, and I'm just trying to hit a few of them. Some people looked at taking the LMS algorithm and extending it very straightforwardly into a direct form IIR filter; some of the early ones I've mentioned here. Some people looked at again taking the least squares, taking the squared error approach, but, instead of just using LMS, preserve the whole equation for the gradient and recurse that. Sam Stearns, Stan White, Nasir Ahmed from the University of New Mexico all looked at various schemes like that. Other people looked at them from a completely different point of view, of ARMA and ARMAX modeling. Ben Friedlander, Martin Morf and a whole bunch of other folks at Stanford and other places looked at that. Rick Johnson, Mike Larimore and I have looked at some other methods. We were using some nonlinear stability theory ideas that were being promoted in the control world. So all sorts of people have looked at it, and what I'm fumbling here for is the next viewgraph. Got it! [VIEWGRAPH #12]

But in fact not much has come of it, and why is that? Well, it turns out that analytically it's a real mess. Those people who have had success with it have found ways to represent the problem as multichannel FIR filtering by various tricks and contrivances, and I'm sure Ben Friedlander will get me later for saying that. The methods that have actually been used, people have had to simplify them rather considerably. A fellow named Paul Thompson at Sandia Labs came up with a method that has been used where you adapt the zeroes and you lock the poles in to sort of follow along on the same radius but back

AST

IIR ADAPTIVE FILTERING

- PROMISE
  - REDUCE COMPUTATION CONSIDERABLY WHEN MODELING RESONANCES OR COMPENSATING FOR SPECTRAL NULLS
  - REDUCED DIMENSIONALITY OF FILTER MIGHT IMPROVE CONVERGENCE RATES OF ADAPTIVE PROCESS
- ANALYTICAL WORK
  - STRAIGHTFORWARD EXTENSION OF LMS -- FEINTUCH, HORVATH
  - RECURSIVE GRADIENT ESTIMATION -- STEARNS, WHITE, AHMED
  - ARMA AND ARMAX METHODS -- MORF, FRIEDLANDER
  - STABILITY-BASED METHODS -- JOHNSON, ET. AL.

VIEWGRAPH #11

B-282-89

AST

IIR ADAPTIVE FILTERING (CONTINUED)

- PRACTICAL SUCCESSES
  - CONSTRAINED DESIGNS
    - POLES LOCKED TO ZERO (THOMPSON)
    - ONLY TWO POLES (CCITT ADPCM ALGORITHM)
  - 1-POLE DESIGNS IN FREQUENCY DOMAIN ADAPTIVE FILTERS
- ANALYTICAL AND PRACTICAL PROBLEMS
  - HIGHER ORDER FILTERS HAVE SLOW AND/OR UNCERTAIN CONVERGENCE
  - CONSTRAINTS/TESTS TO ENSURE STABILITY ARE DIFFICULT AND HARD TO IMPLEMENT
  - EQUALIZATION OF NON-MINIMUM-PHASE CHANNELS REQUIRES UNSTABLE POLE/ZERO CHOICES

VIEWGRAPH #12

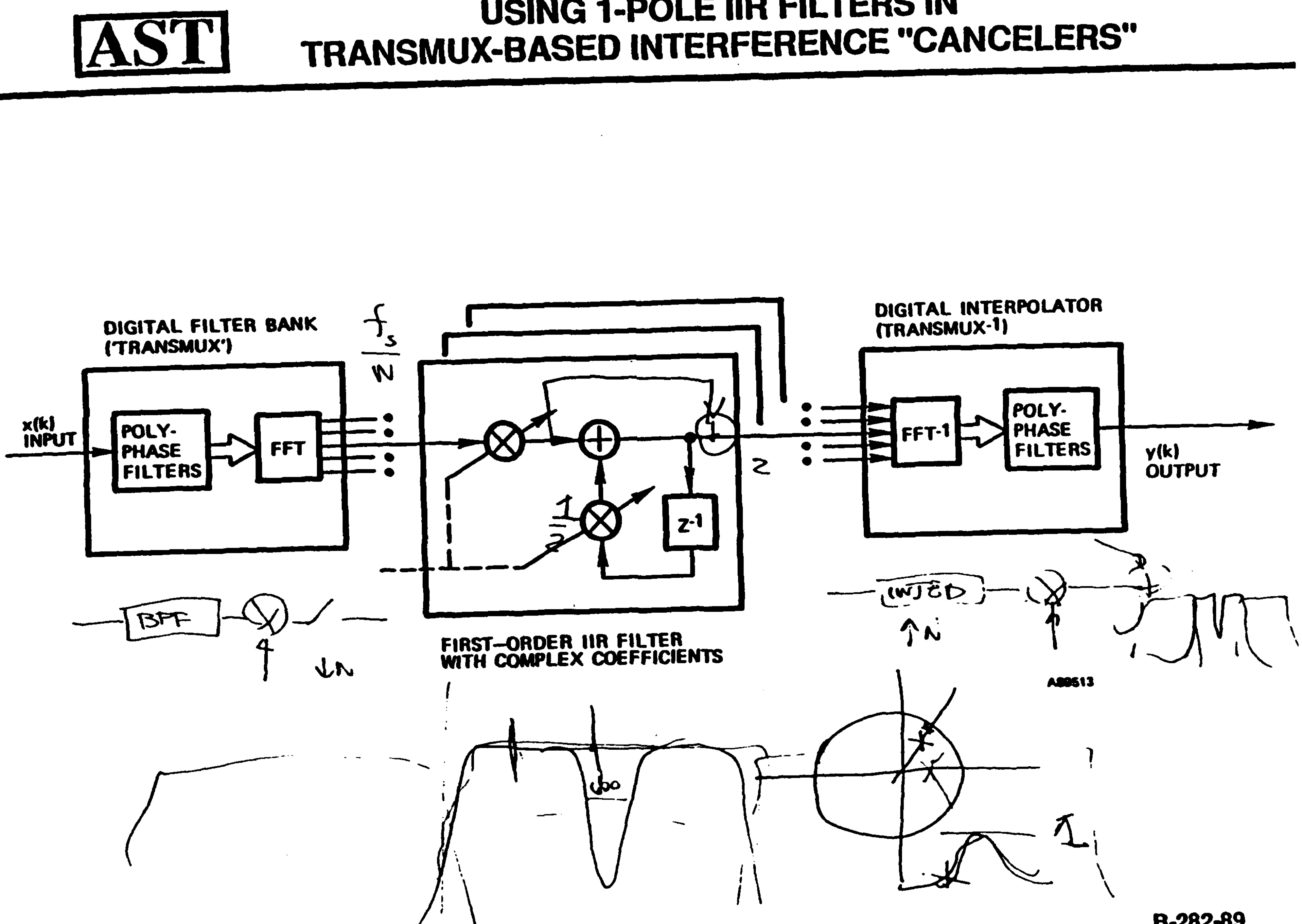
B-282-89

off the unit circle some. There you can simply look at it as an FIR filter and everything's cool. Probably the most practical and common use of it right now is the CCITT ADPCM algorithm, that basically uses linear predictive filtering in order to reduce the dynamic range of voice signals. You can get the VLSI chips to do this and so forth. They use adaptive IIR filtering: 6 forward taps and 2 feedback taps, only 2, and those are badly constrained; they don't want them to do anything bad. They constrain the maximum radius, and all sorts of other things about them. So that's not exactly like you really trust the algorithm, OK? I'll show you a picture in a minute of some other areas where people have melded 1-pole IIR filters, which it turns out you can say a fair amount about analytically and you can prove work, and you can guarantee their stability. I'll show you that picture in a minute.

What have been the problems? Well, there are two analytical problems which I'll talk about first. It turns out that these things are hopelessly hard to analyze by people like me who like nice, simple least squares problems. It's hard to solve for the properties of them; the solutions are hopelessly multimodal. It's hard to come up with schemes that have guaranteed convergence. People have come up with various schemes to do it. As I say, Rick Johnson and Mike Larimore and I did some based on stability theory, where you can in fact guarantee convergence. But you need to know so much about the system that you're trying to model, that you may have well have done something else instead. You need to know enough ahead of time to trick the algorithm into being stable. And there are other schemes for doing it, but all of them have their hard parts. But the real killer has been that if I want to build an equalizer out of this, and there's ever any chance of the channel I'm equalizing being non-minimum phase, even if I've solved all the problems from the convergence and convergence rates and all that sort of thing, the adaptive IIR equalizer wants to put poles outside the unit circle. So IIR filtering appears to be a computational panacea, but then you go look at it and say, "This is crazy, I can't guarantee that my channel is minimum phase," and therefore that my IIR equalizer will remain stable. Realizing that then you don't even consider it any further.

Let me show you one place where it has proved useful ... I may never find the viewgraph again ... there's one on the floor ... aha, I see it [VIEWGRAPH #13] ... is to use it in a situation like this: one of these transmultiplexer-based or channelized interference cancellers. Think of a situation where I've got my signal broken into a thousand bins already – I've already used the big transmultiplexer – but the interference is even narrower than that. Let's consider a practical problem where these bins might be a kilohertz wide. But the interference I'm actually going at is a push-to-talk carrier, and the carrier's come up and he hadn't even gone into modulation yet. Very typical, you punch the button, you click the mike twice, or you punch the carrier and you finally remember what it is that you're going to say, and then you finally start talking. For the first couple of seconds this thing can be a sinusoid. I can view this carrier just as a sinusoid waveform, very narrowband, and I put a single IIR filter in there or a single IIR filter – I didn't show it here, but with a zero on it – and adapt just these two coefficients? I can move that zero and pole right over to where I want that signal cancelled, and instead of knocking out the whole bin, the whole kilohertz wide bin, I can only take out a couple hundred hertz

USING 1-POLE IIR FILTERS IN TRANSMUX-BASED INTERFERENCE "CANCELERS"



VIEWGRAPH #13

B-282-89

or even a few hertz with something like this. People are using this and it does seem like a good idea. You get around all the stability problems because all you have to do is constrain the feedback coefficient to have magnitude < 1, and because you've only got two weights, it's simple to converge. People have been actively looking at schemes like that. But other than that, IIR adaptive filtering is almost stillborn as a field.

CIOFFI: In this case where you're doing this type of thing to do an IIR, the FFT matrix is probably just doing the eigenvalue decomposition, and all you're doing is knocking out the singular values which are small or large. It really isn't an IIR filter; it's just another instance of these maximum likelihood detection type techniques where you form a correlation matrix and knock out the small eigenvalues. So ....

TREICHLER: Precisely.

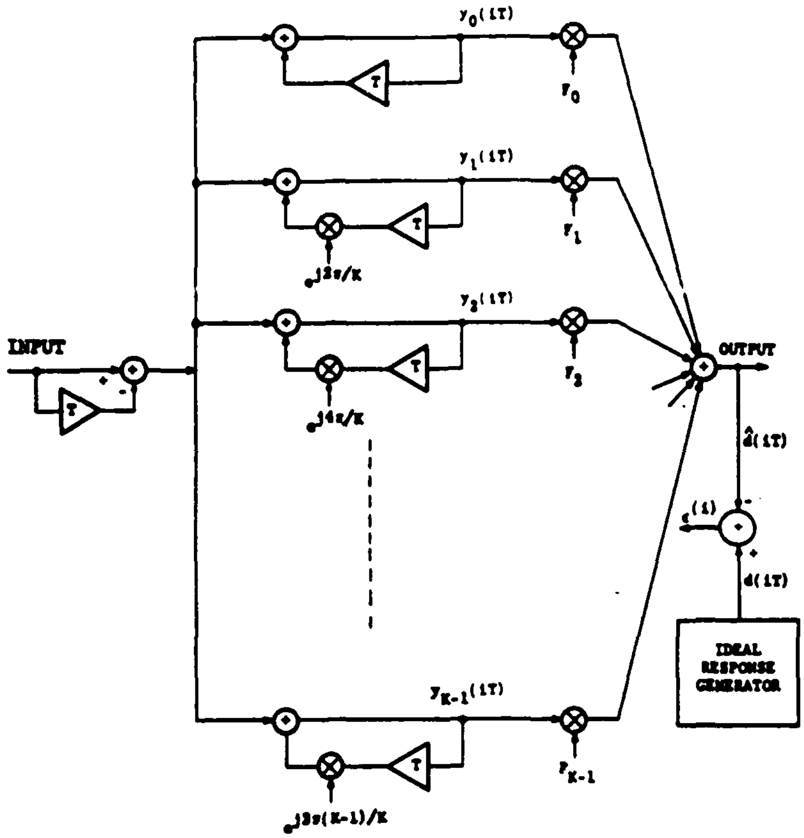
CIOFFI: I would fall in the line that the IIR filter is probably, you know .... There aren't any good applications other than the G.722 thing that you're talking about, which is constrained. Because of the fact that it's just an ill-posed problem. If you take into constraints what's available to you, then maybe you can do it. But generally speaking there really isn't a good solution because it's just ill-posed. It's more of a data compression problem than it is an IIR adaptive filtering problem, because you're trying to reduce the number of parameters that you use to model a system so it falls in the data compression range, and is ill-posed as a filtering problem.

TREICHLER: I don't have any problem with that assessment at all. I'm simply trying to point out that when I go through and say there are no practical applications – in fact if you want to be perverse about it, yeah, I can point to the ADPCM and I can point to this, and there are places where it looks IIR. In fact, the whole point of this is to reduce the paramaterization to a negligible level to where you can use that scheme.

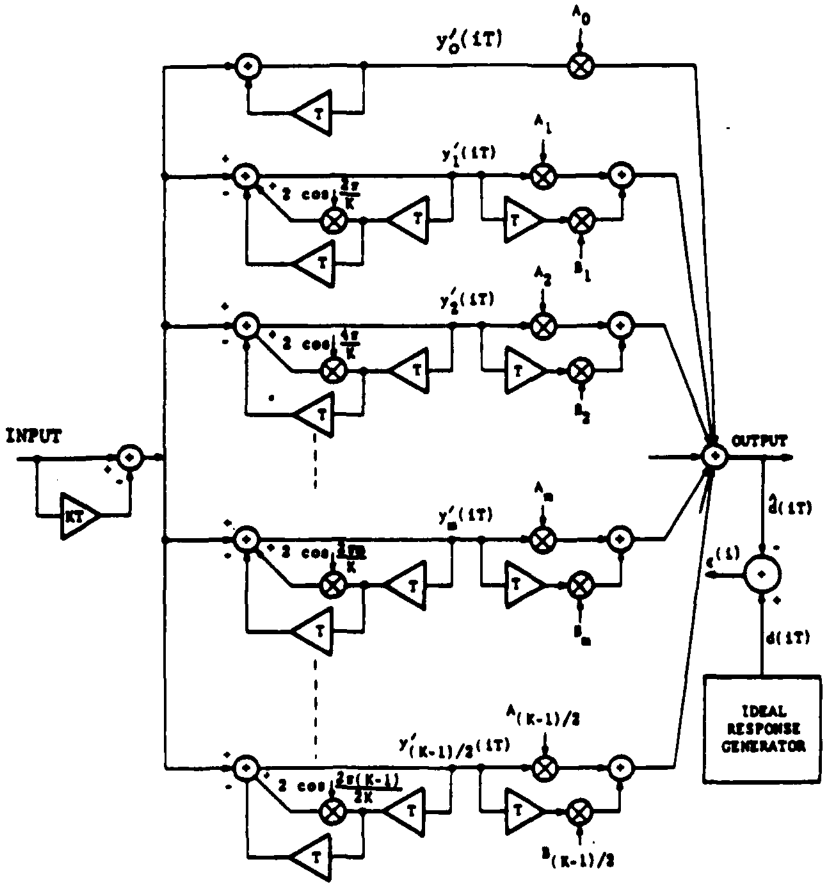
I should point out in passing one of the limitations of the transmultiplexer method is you have to make some guess once of what the right bandwidth is for the bins, OK? And you're always wrong – some signals are bigger, some signals are smaller. Using one-pole IIR filters is a way of keeping your FFT from having the bins too small. You can go after individual very narrowband interferers with simple IIR filters like that.

PROAKIS: I would like to offer an alternative solution to the problem of IIR filtering. There is another structure for an adaptive filter which really has not received sufficient attention. This is a so-called frequency sampling version of an FIR filter, which consists of a comb filter in cascade with a parallel bank of resonators. These coefficients, which can be adjusted, are identical to the DFT coefficients. So they're simply a transformation of the time domain FIR filter parameters. The nice property of these coefficients is that in a sense we obtain orthogonality in the frequency domain through this mechanism, and we can adjust the DFT coefficients independently. If we can put notches in the spectrum, we can also put bumps in the spectrum, depending on what it is that you're trying to do. We can use this for equalization purposes as well as for tone excision purposes. I think that this structure is probably more suited to the kinds of things that you're trying to do with IIR filtering, and I offer this as an alternative.

TREICHLER: I completely agree. But I wanted to share just one other bias with you as we pass through here. In my profes-



A recursive adaptive filter.



A recursive adaptive filter equivalent to the one shown above.

J. Proakis

sional life, I avoid IIR filters like the plague simply because of dynamic range and finite word-length considerations. So if there's ever any FIR way of doing anything – feedforward, pipeline, that sort of thing – that's the way we typically build things. But I completely agree analytically with what John has said about this alternative approach.

AGEE: I have a question for John Treichler. This is Brian Agee. Going back to your original talk, the one you gave in the first half, you talked a lot about the particular problems that the transmultiplexer-based filtering had as far as transport delay and screwing up the poles and everything. How much of a problem occurs if you look at just FFT-based techniques – overlap and add, that sort of thing? Do any of these problems crop up?

TREICHLER: Absolutely. It's only a matter of degree. The first picture that I put just had the FFTs, had a fast convolution/fast correlation method. You end up with a little bit of transport delay there simply because you have to gather up the data into blocks in order to do an FFT, and then you have to read it out the other end. If it actually takes any time to compute anything like the FFTs you end up with more transport delay. Any time you have delay in any of these schemes, it causes you problems with the updating. And while I'm here I also have concerns that the fancier and fancier you make your property restoral method of estimating – looking for a baud tone, looking for a pilot tone in an FDM baseband – the more you have to narrowband to find something, you're going to introduce more delay yet. This is a problem we have to deal with because we'd like to use these fancier schemes. The only reason I focused on the transmultiplexer design is that it sort of represents the limit of the most delay you would throw in, in the filtering alone, much less in the measurement of the error function. It was suggested, what if I window the FFTs? Every time you do that you're introducing more bulk delay and more transport delay, and you get into further trouble with that mechanism.

LLOYD WELCH: Is this a long table discussion or do you want me to sit ...? [LAUGHTER]

VEDAT EYUBOGLU: I want to make a comment that is related to the shaping gain that John Cioffi talked about and blind equalization. It seems like, based on recent research, soon we may be seeing signal constellations that have non-uniform distributions, distributions that are close to Gaussian. On the other hand we know that if the constellation is exactly Gaussian, the output of the channel does not contain any information about the phase of the channel. On the other hand it seems like ... so it would be impossible to do blind equalization if the input constellation was perfectly Gaussian. On the other hand it seems like one could use that to our advantage in applications where we are trying to avoid being detected. That is, try to use a constellation that is, as a distribution, as close to Gaussian as possible. Would you like to comment on that?

TREICHLER: First of all, I can't comment about all blind equalization methods because I haven't looked at them all in equal detail. But it's definitely guaranteed that the dispersion direction – Godard, CMA or whatever – will definitely fail if the signal is Gaussian. As a matter of fact Brian and some guys at our place have actually got formulas in terms of the kurtosis of the constellation that says if it's less than this, it works; and if it's greater than that, like Gaussian, it doesn't. So yes, I would agree with you that Gaussianity in the signal constellation

is going to make it less detectable and more wonderful, but it's going to make life very difficult for people like me.

AGEE: I want to comment on that too. There's a paper by Benveniste in the *Transactions on Automatic Control*, 1980, where he just discusses the general problem of signal identification on Gaussian channels, and he comes to the conclusion, as you've said, that if the signal is Gaussian, if the distribution of the signal is Gaussian, then you're out of luck, and I would agree.

However, I would also say that some of the comments that Seymour Stein made yesterday about how it affects the ability of the communicator to do his job also apply here. As you make the signal harder and harder to intercept, you also make it harder and harder for you yourself to receive. So there's kind of a fundamental paradox that I think you're going to have difficulty overcoming.

A third comment I want to add is that there is one direction to go in if the amplitude of the signal is Gaussian, and that is that the self-coherence restoral techniques are not dependent on the amplitude of the signal. They will work for certain applications – in particular for antenna arrays – if the modulus of the signal is Gaussian, as long as the signal has second order cyclostationarity. That can be gotten around, as Mark Wickert said, by reducing the excess bandwidth of that signal. But if you do get down to lower excess bandwidth, then I think there are directions to go in exploiting higher order cyclostationarity of the signal.

The other question that I'd like to ask is if you want to design a communications system, and you want to make that system low cost, is there a way you can design a system and deliberately build in some known signal properties so that you can make it easy to do blind adaptation at the other end? You know, you can go both ways on that, I guess. How can you make it hard to intercept, how can you make it easy?

EYUBOGLU: Well, it seems like I agree with you that it would be very difficult to achieve a perfect Gaussian distribution. However, it seems on the other hand quite easy to achieve constellations that have non-uniform Gaussian-like distributions. So as you approach the Gaussian distribution, I suspect the detectability is starting to become more and more difficult on non-minimum phase trials.

AGEE: I would agree.

CIOFFI: I'd agree also that the technique – especially, Vedat is asking the question because of the new methods from Codex, called "trellis shaping," which basically will give you that close to the Gaussian type distribution. The input is still an i.i.d. type sequence, so couldn't that be exploited rather than the Gaussian? In other words, you have basically white Gaussian noise as the input to the channel, and what you see on the output is then at least of magnitude characteristic of the channel, which could then be used to, at least, get an initial guess at the equalizer. You wouldn't get any phase information that way about the channel, but the magnitude would still be there.

EYUBOGLU: It's going to be there. The phase information, that will be impossible to extract because the output of the channel is going to be a Gaussian signal, which is going to be perfectly described by its power spectral density. It's not going to contain any information about the phase or the channel.

CIOFFI: I agree with that. You'd only have the magnitude.

TREICHLER: Let me just say that there are still some games you can play.

HALL: John can solve your problem, no matter what! Dr. Peile has a question.

PEILE: Yes, Rob Peile. The subject came back to making an interceptor's life difficult, though Vedat's raised a point. There's another point that occurred to me. If you take the block lattice approach to coding and modulation - you plot a set of points in, let's say, 24 dimensional space - normally you arrange things so the two dimensional projections are nice and structured, so you can treat it as a sequence of complex samples from a nice, structured QAM constellation. If you applied a unity transformation of some sort to the sphere packing or lattice points, you would still have the same Euclidean distance properties, it would be marginally harder to demodulate transmitter/receiver (not much though). But if as an interceptor you were treating it as the sequence of points from two dimensional constellations or complex samples, it would be really unstructured. It wouldn't look like a QAM constellation, and in any particular sample the projections or coordinates would just look like a mess. I think that that would be easy to do, hard to intercept, unless you started treating your samples as multidimensional quantities. People agree with that, any comment?

HALL: Yeah, that makes sense to me. That's another level of complexity that the communicator would add, which you may want to add anyway for coding purposes.

PEILE: I think it would be really easy to implement over, up and above the cost of implementing the thing anyway.

HALL: OK, let's see. Oh, Dr. Gardner - is it short? I have to get on a horse at 1:00 and I want to eat first, so .... [LAUGHTER]

GARDNER: Since the spectral self-coherence restoral approach that Brian has mentioned briefly does seem like a promising alternative to modulus restoral approaches, I just wanted to briefly remark since most of you probably aren't familiar with this idea to show you the analogy to something that you probably are familiar with. This is the adaptive line enhancement technique to separating broadband signals from narrowband signals. Spectral line enhancers simply exploit the fact that narrowband signals have temporal coherence, that is they're correlated with time shifted versions of themselves with relatively large shifts, whereas broadband signals are uncorrelated. So you can process the combination of those signals to restore this temporal coherence at an appropriate time shift to separate the broadband and narrowband signals. The analogy is that in the spectral self-coherence algorithms we're looking at, all cyclostationary signals are correlated with frequency shifted versions of themselves, but only at distinct frequency shifts corresponding to periodicities of cyclostationarity like baud rates and carriers and so on. So by picking the right distinct frequency shift and restoring spectral coherence of that frequency shift you can separate various cyclostationary signals. Thank you.

HALL: OK, I'd like to thank the speakers for a most entertaining session, and being so cooperative in the time limits made my job very easy.

SCHOLTZ: Very briefly, Item No. 1: If the weather gets inclement, if it's too cold, we'll just move the cocktail hour in probably to the Chiricahua Room, which is where the banquet is going to be tonight in the main building.

No. 2: I'd like to talk to the session chairmen for just about one minute up here so Dennis can get on his horse, but I'd like to reach you. I'd like to make sure that all of you come to the Wrap-up Session tomorrow.

After the two talks I'd like to have a critique of the workshop and I'd also like people here, after hearing all this, to say, "Gee, what's left to be done, where should basic research go?" in sort of a broad way without giving away all your favorite secrets.

The last comment is we've had a lot of extra viewgraphs that have come up outside the regular ones; I hope you'll also send copies of those so we can get them into the record.

Thank you all very much, we'll see you at 3:00.

Adaptive Coding

Session Chairman: Robert Peile

M. Vedat Eyuboglu
Coding and Equalization

Allen Levesque
Error Control for the HF Channel

Robert Peile
Variable-Rate Forward-Error Control

Seymour Stein
Systems Perspectives

ROBERT PEILE: Let me start by introducing the panel; Vedat Eyubolgu from Codex, Allen Levesque from GTE, myself from USC, and Seymour Stein from SCPE.

I'm going to take a minute just to comment on the name "Adaptive Coding." This is probably the most ill-defined subject in this workshop, in that adaptive coding means a lot of things to a lot of different people. So I feel the need just to explain, just to talk about the phrase "adaptive coding" for a minute or two. I don't particularly want to define it because I think it's too early, but I want to delineate some areas.

The first area (which also came up in John Cioffi's talk) is that adaptive coding means something in discussion of the bandlimited channel. A lot of people made the point that if you want to continue the gains in the bandlimited channel beyond trellis coded modulation, you do not design a black box containing an equalizer and a black box containing a coding and modulation device; you mix them up a bit. The first adaptive coding talk of this afternoon, Vedat's, is very similar in flavor to John Cioffi's talk. In fact John could have been in this session and Vedat could have been in that session, it doesn't really make much difference. The point is that one can design coding with equalization and adaptive equalization, and people call that adaptive coding. The motivation is clear enough: we're trying to get the last ounce of capacity out of the channel under less than ideal conditions.

Another motivation for adaptive coding – and what a lot of people refer to as adaptive coding – is that people wish to take data communication networks out of their original environment, which was a benevolent office regime, and put them in places that they were not designed for – over radio networks (SATCOM, VHF or HF) – and where there might be a threat of jamming.

Error correction designed for a benevolent regime is not suited to an unbenevolent regime. There's a whole slew of techniques which are appearing and have appeared. You can start off with Chase's code combining as an important concept. Then you talk about hybrid ARQs, Type 1 and Type 2's, and then combinations of Type 1 and Type 2's. They seem to be appearing all over the place; there are a large number of different protocols appearing under different conditions. Under some conditions they give very high throughput and under other conditions they don't. That's a big motive for adaptive coding, to try and get better data communications in tactical military communications. I once pointed out flippantly that mobile communications tend to have mobile channels. Allen Levesque is going to review some of these coding techniques and put them into perspective.

Another point people tend to forget is that, even if you look at what's available off the shelf this moment, there's an awful lot of forward error correction available. Qualcomm offers a chip with 4 or 5 modes of operation; other companies are producing competitive chips. You can go to Cyclotomics and buy a box with dozens of codes in it. If you start to concatenate them, there are literally hundreds of codes you could use. This is the good and the bad news. I have heard a lot of people say that they don't want hundreds of codes, they want to have something that's useful to them. They don't want to train specialists, they just want to have something that helps them. There's an operational requirement to try and automate the process of coding, and that's more in line with my talk.

So these are three areas called "adaptive

Adaptive Coding

Panel:

Vedat Eyuboĝlu	(Codex)
Alan Levesque	(GTE)
Robert Peile	(USC)
Seymour Stein	(SPCE)

Adaptive Coding

Why:

Technical Feasibility
Operational Requirement

Where:

Mobile Communications
Telecommunications
Tactical Military Communications

What Is It?

Error control in real-time

coding". But, of course, "adaptive coding" is part of a bigger thing called "adaptive systems". The final speaker, Seymour Stein, is going to talk generally about which corner of the big picture he sees adaptive coding fitting into. Hopefully, this will lay the foundations for a good discussion. I feel that, as this is an emerging area, the discussion is probably at least as important as any other area, maybe more so. It would be interesting to have some input.

OK, so with that rough overview, I would like to turn over to Vedat ....

VEDAT EYUBOGLU: *Coding and Equalization*

Thank you, Rob. [VIEWGRAPH #1] Since the introduction, by Ungerboeck, of trellis coded modulation schemes around the late '70s and early '80s, we have seen a considerable amount of work trying to extend Ungerboeck's studies. In part these extensions focused on new codes to improve the coding gain. In fact codes were found that offered some modest improvement over what Ungerboeck had described in his original paper. Apart from that, it seems to me, one of the important observations was one made by Forney, who showed that the gain that can be achieved with a coded modulation scheme can be separated into two parts. One is a gain that is due to the fundamental trellis code; the other is a smaller gain that can be achieved by properly choosing the boundary of the signal constellation. Forney called the latter the "shaping gain."

More recently, and as you've seen in John Cioffi's talk earlier today, there is some significant interest in extending coded modulation schemes to channels that not only have Gaussian noise but also introduce linear filtering on the transmitted signal. Here I'm going to introduce a new scheme that can achieve the same kinds of gains that you can on flat channels, on distorted channels, without significantly increasing the complexity.

[VIEWGRAPH #2] First I want to state the problem. This figure is very similar to the one that John showed earlier today. Basically we have a channel that consists of a linear filter and Gaussian noise, white or correlated, and we have a transmitter and a receiver. The transmitter includes an encoder whose output is generated at the baud rate, $\frac{1}{T}$. The sequence that is generated (by the encoder) is passed through a transmit filter that provides pulse shaping. At the output of the transmit filter we have an average power constraint. In the receiver, we have a receive filter whose output is again sampled at the baud rate, and a decoder that tries to estimate the bits that were transmitted. The main assumption here is that Gaussian noise and linear distortion are the dominant impairments that we are trying to deal with. This is typically the situation on voice band modems, for example.

I'm going to restrict my attention to modulation schemes that are single carrier. I want to do this because most modems use single-carrier schemes and therefore there is much more experience in them. Also, as John explained earlier today, single-carrier systems can achieve the same amounts of gain that can be achieved with multi-carrier systems.

The problem that we are trying to solve here is to design an encoder-decoder pair, as well as choose these (transmit and receive) filters, so that we can maximize the bit rate when the probability of error and the complexity are restricted to lie below certain limits. As it turns out, we don't have to redesign the encoder-decoder pair; we can use the same codes that have been designed for white Gaussian noise channels. As I'm going

COMBINED CODING AND EQUALIZATION FOR LINEAR DISTORTED CHANNELS USING TRELLIS PRECODING

BY

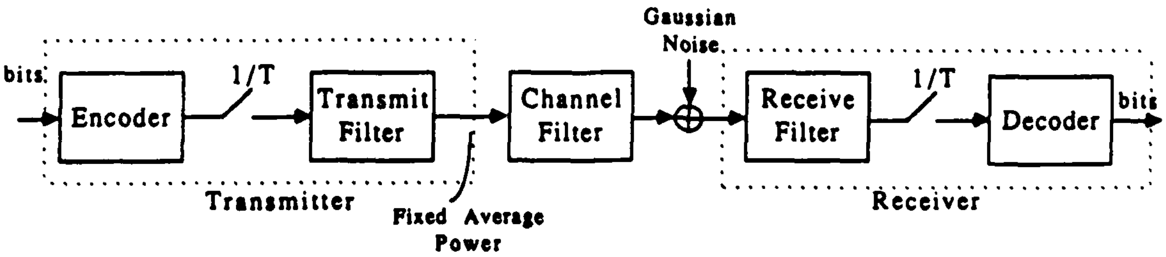
M. VEDAT EYUBOGLU

CODEX CORPORATION
MOTOROLA

VIEWGRAPH #1

CSI WS-5/16/89

PROBLEM STATEMENT



Assumptions: Gaussian noise and linear distortion are the dominant impairments; Modulation: Single-carrier QAM.

Problem: Choose the encoder/decoder and the filters to maximize bit rate, while probability of error and complexity are kept below certain limits.

VIEWGRAPH #2

CSI WS-5/16/89

to explain, with proper precoding operations, these codes can be used on distorted channels as well.

[VIEWGRAPH #3] First, I'm going to introduce a classification of applications. The scheme that I'm going to describe and the multi-carrier one that John described require knowledge of the channel at the transmitter. If such knowledge is available, and if in addition the receiver also has knowledge about the channel, then we can have joint transmitter-receiver equalization. This can be done in applications that are point-to-point, and where the channel is time invariant or very slowly varying.

There will be some applications where equalization cannot be done at the transmitter. This includes, for example, broadcast channels where you may have one transmitter simultaneously communicating with two receivers. In this case, since effectively the transmitter is seeing two different channels, it wouldn't know which channel to equalize. Another example where transmitter equalization wouldn't be possible is a simplex channel, where the measured channel information cannot be sent back to the transmitter.

There could be a third class, which I call "pure transmitter equalization." It may be desirable to do equalization entirely in the transmitter. An example of this is a polling system where one would like to achieve rapid inbound training. If all the equalization is done in the transmitter, it's not necessary to train the receiver in every poll. This also avoids any address recognition requirements that are necessary for coefficient-storage schemes that use receiver equalization.

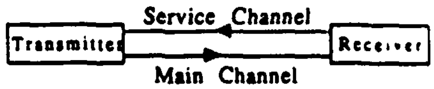
[VIEWGRAPH #4] This viewgraph shows a brief summary of known equalization schemes for coded systems, some of which have been discovered in the last couple of years. Again, I'm focusing on single-carrier systems. You can see that the first three schemes are the ones that John Proakis described: linear equalization, decision feedback equalization, and maximum likelihood sequence estimation. All three could be used in a coded system.

Linear equalization is the simplest. It can be used either as a joint equalization scheme where part of the equalization is done in the transmitter and part in the receiver, or done entirely in either the transmitter or the receiver. There is one difficulty when one uses linear equalization in conjunction with a code that is designed for a white Gaussian noise channel. The noise, after it passes through this receive filter, becomes correlated at the input of the decoder, and that can create a mismatch, and degrade the performance. But there is a simple way of getting around that by putting an interleaver in the transmitter, and a deinterleaver in the receiver before the decoder.

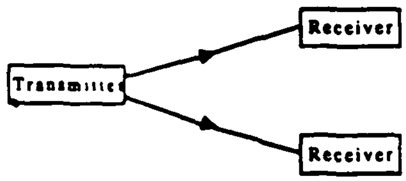
The second scheme is decision feedback equalization. As I'm going to illustrate, it plays a central role in combined coded modulation and equalization. It turns out that the performance of the ideal DFE is the performance that we should be trying to achieve, and we can in fact approach capacity with an ideal DFE. However, in a coded system, DFE cannot be applied in a straightforward manner, because the decoder in a coded system has to delay decisions, whereas a decision feedback equalizer requires decisions for feedback immediately. We proposed, some time ago, a scheme that uses interleaving in a somewhat different way to make delayed decisions available for decision feedback equalization. It's used with a predictive-form decision feedback equalizer and it works well provided

A CLASSIFICATION

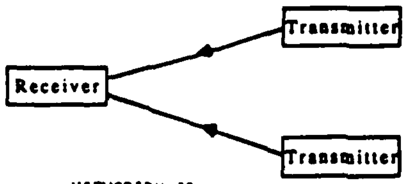
A. JOINT TRANSMITTER/RECEIVER EQUALIZATION: Transmitter and receiver both have knowledge about the channel. (Point-to-point and Time-invariant Channels)



B. RECEIVER EQUALIZATION: (Broadcast Channels; also Simplex or Time-Varying Channels)



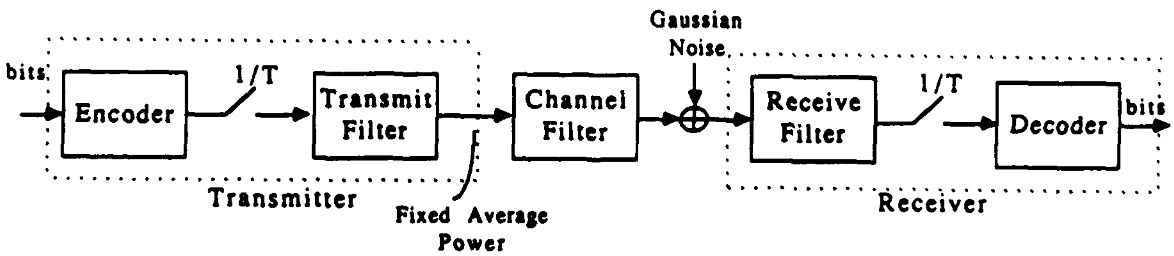
C. TRANSMITTER EQUALIZATION: (Polling System with rapid inbound training)



VIEWGRAPH #3

CSI WS-5/16/89

EQUALIZATION ALGORITHMS FOR CODED SINGLE-CARRIER QAM



- Linear Equalization with Interleaving (Joint)
- Decision-Feedback Equalization with Interleaving (Joint)
- Maximum-Likelihood Sequence Estimation (Joint or Receiver)
  - Reduced-State Sequence Estimation (RSSE)
  - M-Algorithm
- Generalized Precoding
- Trellis Precoding (Joint or Transmitter)

VIEWGRAPH #4

CSI WS-5/16/89

that the feedback filter is not very long and provided the system can tolerate long delays.

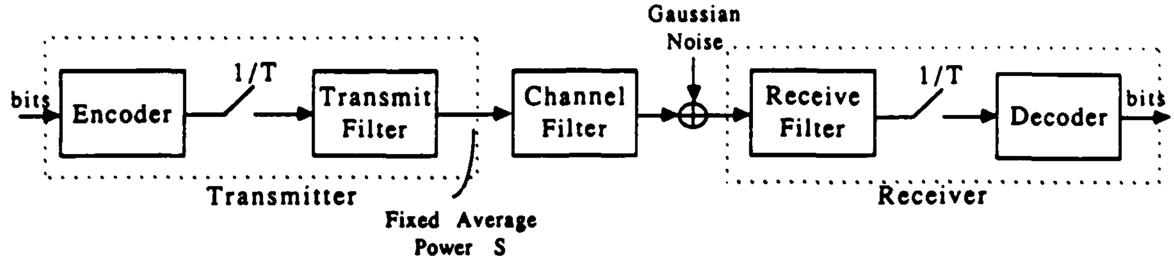
The third scheme is maximum likelihood sequence estimation. This is the optimum receiver technique. Take any of the Ungerboeck codes, pass it through some channel filter, and then try to design your decoder in such a way that you not only decode the code, but at the same time resolve the intersymbol interference. It's well known that the complexity of maximum likelihood sequence estimation can be very high. In fact it increases with the complexity of the code, the length of the channel impulse response that is seen by the maximum likelihood sequence estimator, as well as the number of bits transmitted per symbol, which is typically large in trellis coded systems. We proposed a scheme, called RSSE, that reduces the complexity of MLSE in a structured and systematic way. It uses set partitioning principles to construct reduced-state trellises that can then be searched with the Viterbi algorithm. There is an alternative way of achieving reduced complexity sequence estimation, and that is the M-algorithm which is somewhat less structured than RSSE. Maximum likelihood sequence estimation is of course a receiver technique and these reduced complexity versions could be very attractive if in fact we are not allowed to use any equalization in the transmitter.

[VIEWGRAPH #5] Now I'm going to turn to the case where I'm allowed to use transmitter equalization. Here the transmitter has information about the channel. I'm going to describe a scheme called "trellis precoding." This scheme is also sufficient to approach capacity and its complexity is not high; it appears to be a very practical system. Before I describe it, I want to first reduce this model that I showed (in the upper figure) to a discrete equivalent form. Here I choose the transmit filter as simply a flat rectangular filter. The receive filter is chosen such that it provides whitening for the noise sequence, and at the same time it provides a causal response at this point (decoder input); I'm going to call this minimum-phase response $h(D)$. This is the same kind of filtering structure that you would use in a decision feedback equalizer.

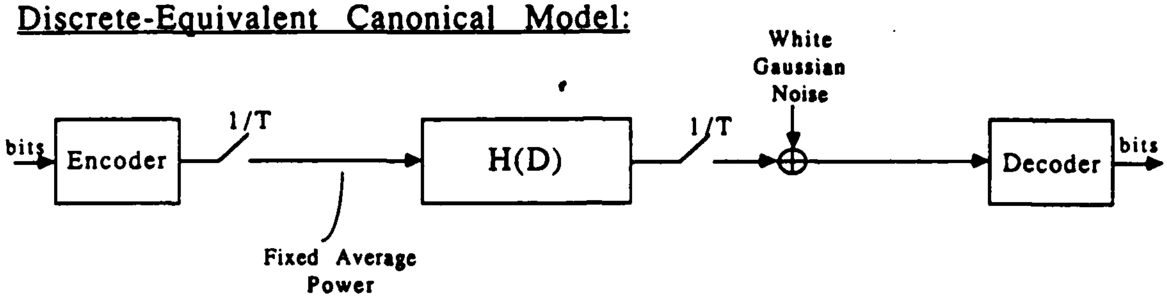
[VIEWGRAPH #6] So now the discrete-time channel that we have consists of a causal minimum phase response, $h(D)$, and a white Gaussian sequence $n(D)$. We encode bits into a sequence, $e(D)$, and these are passed through the channel filter and received at the input of the decoder which tries to estimate the bits. First, I'm going to illustrate the significance of the decision feedback equalizer from the channel capacity point of view. I'm going to focus on the zero-forcing case because my main interest is at high signal-to-noise ratios. There are strong indications that one could reach similar conclusions at low signal-to-noise ratios by using the mean squared error criterion.

In the ideal zero forcing decision feedback equalizer, we use the knowledge of the transmitted symbols, in this case $e(D)$, and pass them through a feedback filter to cancel the tail of the impulse response $h(D)$. Once you do that, you are left with a channel that has essentially one term, the first term of the impulse response, h_0. This is a discrete-time ideal channel with white noise. Interestingly you can show – and this is due to Price, shown in a paper that was published in 1972 which is, I believe, of very high importance and has been overlooked until recently – that at moderate to high signal-to-noise ratios when you compute the capacity of the channel, this is what you find ($C = log_2 S|h_0|^2/N_0$). Here, S

(JOINT) EQUALIZATION WITH TRELLIS PRECODING



Discrete-Equivalent Canonical Model:



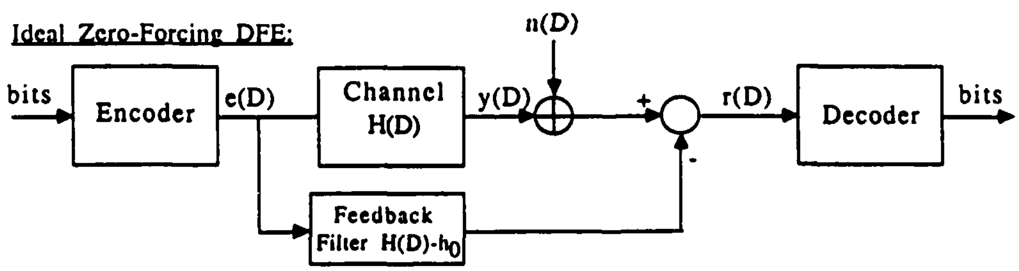
- H(D) is minimum-phase.

VIEWGRAPH #5

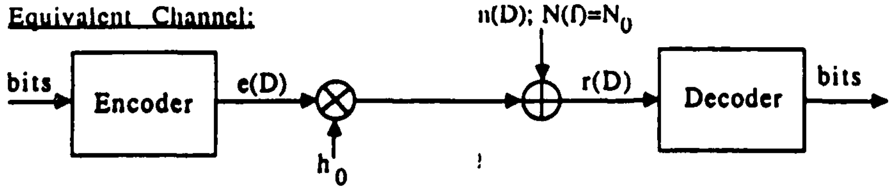
CSI WS-5/16/89

CHANNEL CAPACITY vs. IDEAL DFE

Ideal Zero-Forcing DFE:



Equivalent Channel:



Channel Capacity: At moderate-to-high SNRs, the channel capacity is given by $C = \log_2 (S|h_0|^2/N_0)$ (Price 1972) ; to approach capacity the transmitted sequence must have a Gaussian distribution (Shape gain). If a coding scheme can approach capacity on the flat channel, then the same code can approach capacity on the distorted channel, provided that the performance of ideal DFE can be achieved.

VIEWGRAPH #6

CSI WS-5/16/89

is the maximum power that is available at the output of the encoder, the power constraint, and N_0 is the spectral density of the noise. This is saying that the channel capacity depends on the impulse response of the channel only through the first coefficient, h_0. It doesn't depend on the tail of the impulse response. This is surprising because it's been argued in the past that one advantage of maximum likelihood sequence estimation was that it exploited the energy in the tail of the impulse response, and arguments were made that a DFE is in fact a suboptimum structure. However, when you approach capacity this doesn't hold anymore. You can show that a zero forcing DFE is in fact sufficient to approach capacity. Furthermore, the channel that the zero forcing DFE constructs or generates is an ideal channel. There is just scaling and white noise. Therefore we can use the codes that were designed for the white Gaussian noise channel. In fact if one had a code that could approach capacity on the white Gaussian noise channel, you could take the same code and use it here (on the distorted channel) and it should also approach capacity. The problem, of course, is that the ideal DFE performance cannot be achieved because of the decision delays.

One other comment: In order to approach capacity, the transmitted symbols must have a Gaussian distribution, or, when you look at it in higher dimensions, the boundary of the signal constellation in higher dimensions must be a hypersphere. So an effective scheme that can approach capacity must be able to achieve this shaping gain, it must be able to preserve the coding gain of the encoder that we get in the ideal channel, and furthermore must be able to achieve the ideal zero forcing DFE's SNR performance. The scheme that I'm going to describe (trellis precoding) can achieve all of these.

[VIEWGRAPH #7] This is the only slide that I'm going to show describing the principle of the scheme. The upper part is the transmitter, the lower part is the receiver. It may look complicated but actually from an implementation point of view it's really not. We have two trellis codes. One trellis code is used in the standard way, like Ungerboeck did, to achieve large minimum distance. But there is a second trellis code that we use which is much simpler; it doesn't have to be a complex code; for example, a 4-state Ungerboeck code could be sufficient. That code is going to be used for precoding to achieve the performance of the ideal DFE, and furthermore to achieve shaping to give us a Gaussian-like distribution to conserve average power.

The encoder collects bits in groups, as in Ungerboeck coding, and uses a trellis code. This, for example, could be one of the 1- or 2-dimensional Ungerboeck codes, or it could be a 4- or 8-dimensional Wei code, or any code that is based on binary lattices. This code adds redundancy and we obtain a new set of bits. These are mapped with a simple mapping operation into a sequence of complex symbols that I'm going call $x(D)$. This mapping is, from a complexity point of view, no different from the mapping that would be used in an Ungerboeck code. But it has to satisfy certain requirements. The region where $x(D)$ has to lie is important.

Now this sequence ($x(D)$) is passed through a filter, $g(D)$, which is the inverse of the channel $h(D)$. That gives us the sequence $x\prime(D)$. Then we pass $x\prime(D)$ through a decoder for the second code. So we are using a decoder in the transmitter, a decoder for the second code, the code that is used for shaping and precoding. It's a simple code, therefore

this is a simple decoder. This decoder tries to select a code sequence, $c(D)$, such that the mean square error between $c(D)g(D)$ and the input sequence, $x'(D)$, is minimized. I'm going to say a little more on how this can be done, but let's continue and look at the output. $e(D)$ is then a sequence that is the input $x(D)$ minus the code sequence $c(D)$ that is chosen by this decoder times the channel inverse $g(D)$. After you pass through the channel filter $h(D)$, $g(D)$ disappears, and we are left with $x(D) - c(D)$, which we call $y(D)$. So you can see that the sequence that we are going to receive is the input sequence $x(D)$, except it's translated by a sequence $-c(D)$ from the shaping code. Then white Gaussian noise is added. There is one requirement here which can be satisfied easily; this second code must be a subcode of the first trellis code. Then one can show that $y(D)$ also belongs to the first code. Therefore we can simply use the same kind of decoder that we would use to decode the first code to obtain an estimate of the sequence $y(D)$, and I'm going to call that $\hat{y}(D)$. The complexity of this decoder is not higher than the complexity of the decoder that you would use on an ideal channel.

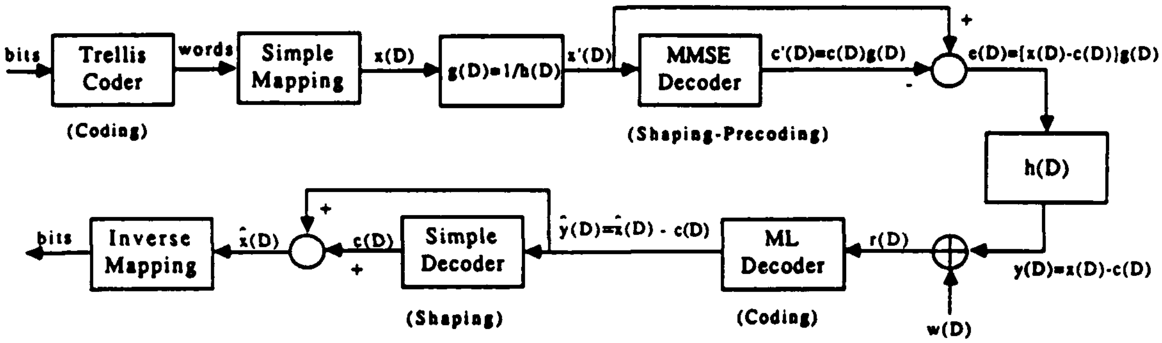
Assuming that you haven't made any errors, you have recovered $x(D)-c(D)$. It turns out, because of the region in which this $x(D)$ was chosen, a trivial decoder can extract the sequence $x(D)$, even though we don't know what $c(D)$ is. Then a simple inverse mapping gives us the bits.

This scheme can achieve the full gain of the trellis code, it can achieve the full gain of an ideal DFE, and it can achieve nearly the shaping gain that can be achieved with a scheme called "trellis shaping" introduced by Dave Forney. [VIEWGRAPH #10] It has in fact trellis shaping as well as the Tomlinson filtering [VIEWGRAPH #9] as special cases. If the channel was an ideal channel, the scheme becomes trellis shaping. This is a scheme that gives you only the shaping gain; it has no precoding gain. On the other hand, if you chose the second code to be just an integer lattice and made your decoder decisions with no delay, then this becomes Tomlinson filtering, which was invented back in 1971.

[VIEWGRAPH #7] The decoder used in precoding is similar to a maximum likelihood sequence estimator. It's a decoder for a filtered code. So its complexity can be high. In fact the optimum decoder cannot be a trellis decoder, it has to be a tree decoder. However, one could use reduced complexity decoding schemes like RSSE to drastically reduce its complexity. We showed that one can use a trellis search based on a trellis that is only as complex as the second code that you're using. If you are using a 4-state code, you can get shaping gains that are often close to the shaping gains that you get with trellis shaping.

[VIEWGRAPH #8] Some implementation issues. As I said, this is a practical system. Of course the transmitter has to have knowledge about the channel, and this can be done by using a training procedure before data transmission starts. One could simply send proper training symbols from the transmitter and adaptively learn a linear receive filter. This receive filter would in fact be the feedforward part of a decision feedback equalizer, which can be broken into two parts: a linear equalizer, whose output is sampled at the baud rate, followed by the filter $h(D)$. One could learn the linear equalizer first and then learn $h(D)$, which is actually a prediction error filter for the residual noise that we see at the output of the linear equalizer. The information about $h(D)$ could be sent back to the transmitter, and it can then be held fixed

TRELLIS PRECODING



- A powerful trellis code is used for coding, another simpler trellis code is used for shaping and precoding. Tomlinson filtering and trellis shaping are special cases.

- Achieves full gain of the code, full gain of ideal DFE and nearly the same shape gain as trellis shaping.

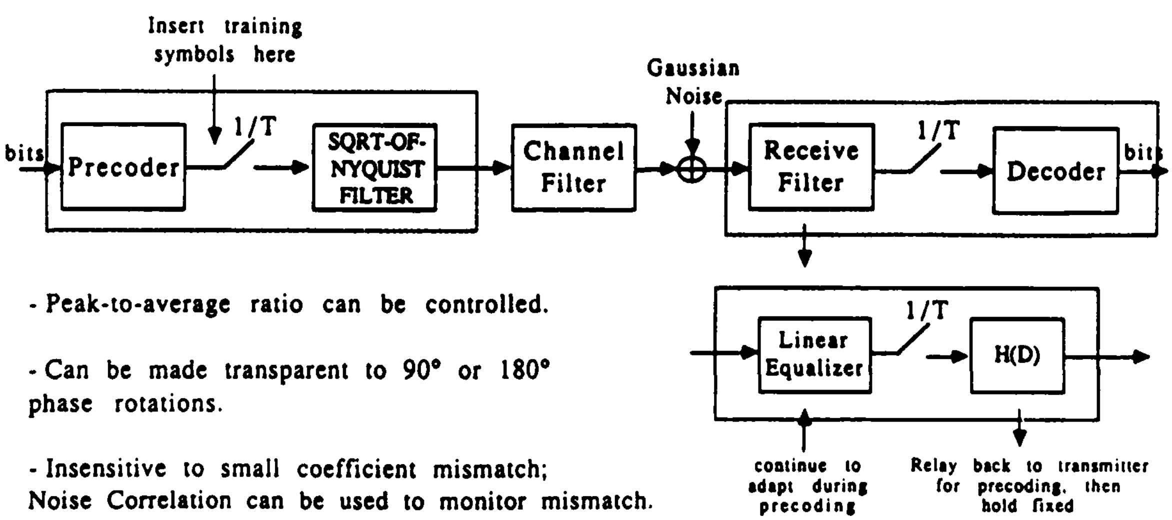
- Shaping and precoding involves decoding a filtered trellis code; it can be implemented using reduced-state sequence estimation techniques.

- Experiments show that for the 4-state 2D Ungerboeck code parallel decision feedback decoding can provide nearly the same shape gain as trellis shaping.

VIEWGRAPH #7

CSI WS-5/16/89

SYSTEM IMPLEMENTATION OF TRELLIS PRECODING



- Peak-to-average ratio can be controlled.

- Can be made transparent to 90° or 180° phase rotations.

- Insensitive to small coefficient mismatch; Noise Correlation can be used to monitor mismatch.

- MSE criterion can be used to improve error rate at low SNR's.

- Increased sensitivity to phase jitter

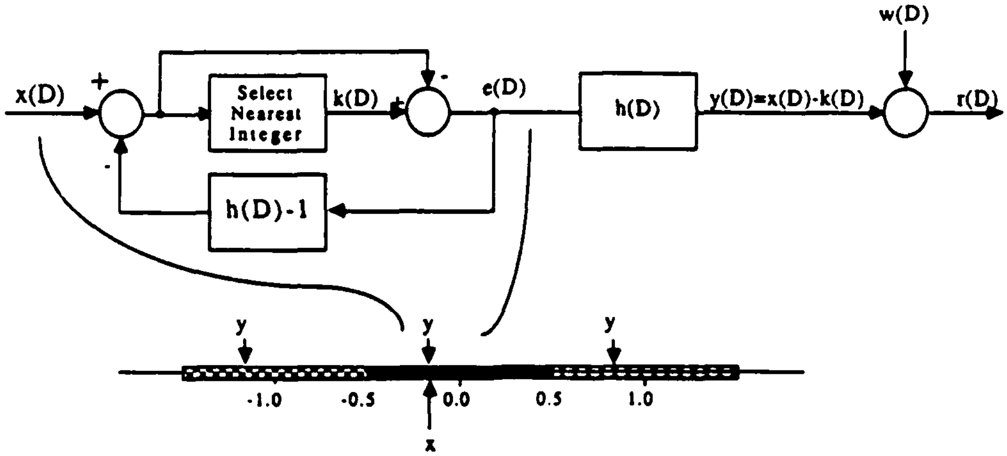
- The baud rate (1/T) can be optimized using off-line channel probing.

VIEWGRAPH #8

CSI WS-5/16/89

TOMLINSON PRECODING

- M. Tomlinson (1971), H. Miyakawa and H. Harashima (1969)



- At moderate-to-high bits/symbol, Tomlinson filtering can achieve the performance of ideal DFE.

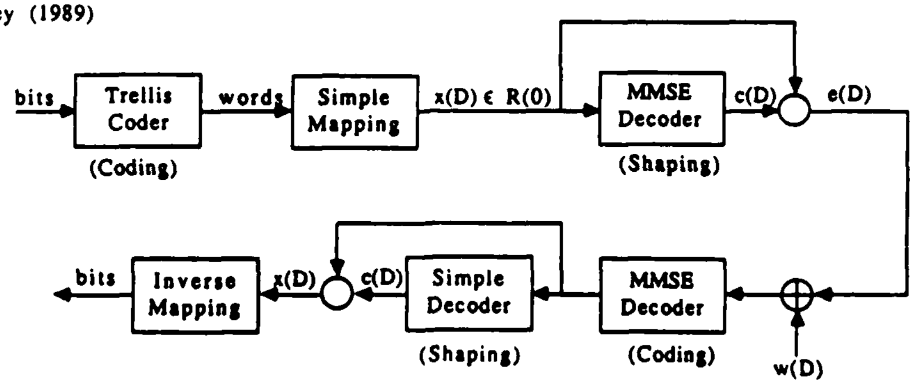
- This has no coding gain and no shaping gain.

VIEWGRAPH #9

CSI WS-5/16/89

TRELLIS SHAPING

- Forney (1989)



- The transmitted sequence lies inside the Voronoi region of a trellis code.

- Simple 4-state shaping codes can provide 1 dB of shape gain.

- The coding gain of the trellis code is preserved.

- Error propagation is avoided using syndrome decoders.

VIEWGRAPH #10

CSI WS-5/16/89

during data transmission whereas the linear equalizer could continue to adapt.

Some system considerations: The peak-to-average ratio of the system, if you don't do anything about it, could be high, because you are getting shaping gain, your constellation is becoming Gaussian-like. However, we found that we can put constraints, simple constraints, on the decoder in the transmitter to control the peak-to-average ratio and bring it close to what we would get on an ideal channel. We sacrificed little in shaping gain. The scheme can also be made transparent to 90° or 180° phase rotations. So if you use a rotationally invariant code like the Wei code that is used in the V.32 modem standard, then, provided you do the right things in the shaping and precoding process, the whole thing can be transparent to 90° or 180° phase rotations.

We don't think the scheme would be sensitive to small coefficient mismatch. As I said, at low signal-to-noise ratios the mean squared error criterion can be used; in fact, in practice that's what you would probably be using since it's easier to adapt.

Finally, so far, I've assumed that the baud rate was fixed. But to approach capacity on an arbitrary channel, one would have to optimize the baud rate as well, and that could be done using an off-line channel probing scheme.

I have one more slide on transmitter equalization, but I'm going to stop. If you have any questions I can show that later on.

JOHN CIOFFI: On your trellis precoding node, the one slide that you had up, I wonder if you could put that back up and I'll quickly ask the question. You have there $g(D) = \frac{1}{h(D)}$ just preceding the minimum mean square error decoder. A typical approach with Tomlinson filtering and such, the equivalent of the decoder is actually inside the feedback loop and the g equals 1 over ... for reasons of stability, one thing, and for instance if you have a notch, at DC, or something like that in the channel, $h(D)$, you're basically increasing your sensitivity to round-off error and that. So is that actually how you're implementing it, or are you somehow sliding the decoder inside that filtering box to put a dynamic range limit on the outputs?

EYUBOGLU: This doesn't have to be the way the scheme is implemented but I thought that this is probably the easiest to understand. You can represent this whole thing in different forms.

ALLEN LEVESQUE: *Error Control for the HF Channel*

The title that you see here on the view-graph is somewhat different from what was printed in the agenda. [TITLE SHEET] I seemed to have come up with a third variation of the title that combines equalization and coding. It seems to be a popular form for titles here ... maybe there's a message there about what subjects we should be pursuing in the future. When Rob was organizing the session and asked us to prepare comments on various aspects of adaptive coding, he also made the comment that we should add some remarks about what we saw as future worthwhile areas of research that should be pursued, items of particular importance that should be given some emphasis. I took a cue from that and, rather than talking strictly about adaptive coding, what I'm going to talk about is a somewhat broader idea of adaptation that would use capability that is available or is going to become available in various adaptive equalizer modems for radio channels, and to use that capability in conjunction with coding to try to do a more powerful, certainly a more flexible, job of adaptation to

COMBINED ADAPTIVE EQUALIZATION AND ADAPTIVE CODING FOR HF LINKS AND NETWORKS

A. H. LEVESQUE

CSI WORKSHOP ON
ADVANCED COMMUNICATION PROCESSING TECHNIQUES

16 MAY 1989

GTE GOVERNMENT SYSTEMS CORPORATION
ELECTRONIC DEFENSE COMMUNICATIONS DIRECTORATE
100 FIRST AVENUE
WALTHAM, MASSACHUSETTS 02254
TEL. (617) 466-3729

TITLE SHEET

OUTLINE

- HF CHANNEL CHARACTERISTICS
- NETWORKING FOR MILITARY APPLICATIONS
- ADAPTIVE CODING TECHNIQUES
- DEVELOPMENT OF SINGLE-TONE ADAPTIVELY EQUALIZED MODEMS
- CONCEPTS FOR COMBINED ADAPTATION OF MODEM AND CODING PARAMETERS

SHEET #2

both links and eventually in networks. The talk is going to concentrate specifically on HF channels.

HF is certainly of importance to the Army, one of our sponsors here, and other branches of the military. It's a medium for which we in our own company, and other companies as well, have done a lot of work in developing a variety of modems, both conventional and frequency-hopped, though I'm not going to get into the issues of AJ today. Besides that, HF seems to be almost everyone's favorite multipath fading channel. It's so troublesome that we can continue to come up with clever tricks that we can develop and test, and if they don't work as well as we'd like we can always shake our heads and say, "Oh, well, you know what HF is like, but next year we'll have a better idea."

[SHEET #2] I'll say just a little bit about HF channel characteristics, but I really won't have to say much because this was covered very well in Phil Bello's talk yesterday and the discussions that went with that. I'd like to say just a little bit about networking and about the developments that have been going on in applying adaptive equalization to radio channels. We've heard a lot about adaptive equalization this morning and again with Vedat's talk on another area. There are some special issues that arise in radio channels that need to be talked about. And I'll talk about some comments for combining these forms of adaptation.

[SHEET #3] We all know that HF is a notorious multipath fading channel. In his talk yesterday, Phil Bello emphasized something that I really should have included in the slide in probably big, bold letters, and that is that HF is a nonstationary channel. It's non-Gaussian but perhaps the more troublesome aspect is the fact that it can be very non-stationary. Typically HF channels are talked about in kind of a rough classification; that is, their fading characteristics as being separable into short-term, well-known, Rayleigh fading, complex-Gaussian fading, and long-term fading which is approximately log-normal. This is a rough characterization, but it's useful to think about the channel that way.

It would be very desirable in trying to design systems for HF channels and eventually HF networks to be able to go to a book or a library or database that would provide us the statistics on the various combinations of channel characteristics. By this I mean signal strength, and multipath, and Doppler, and I'm not going to go off into these in detail because Seymour Stein may comment on some of these same issues. In fact, if anyone wants a good, quick learn on HF, I certainly recommend the article that Seymour wrote in the February 1987 issue of the *JSAC*. I'll mention some of the points that he raised. It would be nice to be able to say that if the signal strength is such and such, then we will also know that the multipath spread is as follows, and would know better how to adapt the modem or make a selection of modem or coding for the channel. Unfortunately that simply does not exist. There are broad statistics about the characteristics of the channels in these various categories, but that ideal source of multivariate statistics simply does not exist. So with all of this it would be very desirable in designing HF links and networks to be able to provide means of adaptation to both short-term and long-term variations, and being able to do this with the lack of the kind of data that we would really like to have to design the links optimally.

[SHEET #4] I want to comment on HF networking because people in many organizations are coming to the realization that the

HF CHANNEL CHARACTERISTICS

- MULTIPATH, FADING, NONSTATIONARITY
- SHORT-TERM (RAYLEIGH) AND LONG-TERM (LOG-NORMAL) VARIATIONS
- LACK OF MULTI-VARIATE STATISTICS FOR HF
- DESIRABILITY OF ADAPTATION TO BOTH SHORT- AND LONG-TERM VARIATIONS
- (REF.: S. STEIN, JSAC FEB87, *pp.* 68-89)

SHEET #3

CSI WORKSHOP · 5/16/89 (4)

HF NETWORKING

- CONNECTIVITY WILL RELY ON MULTI-LINK ROUTING
- USEFUL TO "NEGOTIATE" BEST LINK THROUGHPUT RATE TO ESTABLISH END-TO-END CONNECTIVITY
- ADAPTIVE EQUALIZATION OR ADAPTIVE CODING ALONE MAY NOT PROVIDE SUFFICIENT FLEXIBILITY FOR THROUGHPUT OPTIMIZATION

SHEET #4

CSI WORKSHOP · 5/16/89 (4)

way to make full use of the HF medium is going to be through networking. There's certainly work going on in the services on various aspects of networking. There's work going on inn techniques for automatic link establishment which is a vital ingredient in getting a network up and making it work. But just to say a couple of things ... certainly connectivity in HF has to rely on multilink routing. A very simple example of this could be simply sketched here by noting that if we want to get from point A to point B at a given time of day or night there – at that distance, even if it might be a relatively short distance of a few hundred miles – there may be no HF frequency available to make that connection. However, we might be able to get from point A to point C through a channel on a much longer link and then back to point B. That, in a snapshot, is the kind of thing that effective and efficient HF networking will enable us to do. When you look at dependence upon multilink routing like this, one of the issues that comes up is trying to match data rates on the links that you're putting together to try to make the circuit. It would be useful to have the capability to, for example, negotiate the best link throughput on two or three links in cascade in order to be able to establish the end-to-end connectivity. The real point I want to get to here is that adaptive coding individually or adaptive equalization by itself may simply not provide the kind of flexibility that we would like to have for throughput optimization or for operating effectively in networks.

[SHEET #5] Rob was kind enough to make a few remarks about adaptive coding to save me a few words. I'll get down to the point here on my own view. That is that while the name has been applied to a number of coding techniques, the only area of techniques that really makes sense to me to have the label "adaptive" is various forms of hybrid ARQ. I've written this here just to define a few terms for the sake of some people that might not be familiar with these terms, and also to try to make a start toward a definition. I can say what I don't regard as "adaptive" and that's the simplest forms of ARQ with simple repetition. By the way, the term hybrid refers to the incorporation of error correction with error detection. Simple ARQ you simply block the data, attach error detection parities to every block and repeat the blocks that don't make it. Hybrid incorporates forward error correction and that's well established to be the right way to do it. Simply repeating is not what I call adaptive any more than redialing the telephone when you don't get your party is adaptive. I think the term adaptive is properly applied when we at least begin to talk about saving and combining code blocks that were previously transmitted with newly transmitted code block. Simple versions of this have been looked at; for example, majority voting across copies.

Rob referred to code combining, defined by Chase, in which simply speaking you send a forward-error-correction code; if it doesn't decode properly you send another copy and you combine these on a steady signal Gauss noise channel. You would do, in fact, optimal coherent combining. Without saying a lot about it if you implement this as it's properly defined with maximum likelihood decoding, it's a very efficient scheme because on a steady signal Gauss noise channel, the communication efficiency as measured in E_b/N_0 stays exactly the same regardless of how many copies you have to combine.

There are other refinements of hybrid ARQ that have been looked at. (See Bibliography.) Type-II hybrid ARQ is an efficient adaptive

HYBRID FEC/ARQ CODING TECHNIQUES

NONADAPTIVE
- HYBRID FEC/ARQ WITH SIMPLE REPETITION OF ENTIRE CODE BLOCKS ("TYPE-I")

ADAPTIVE
- HYBRID FEC/ARQ WITH COMBINING OF PREVIOUS COPIES (SINDHU; LAU & LEUNG; BENELLI; CHASE; KALLEL & HACCOUN, *etc*)
- HYBRID FEC/ARQ WITH INCREMENTALLY TRANSMITTED REDUNDANCY, "TYPE-II" (MANDELBAUM; METZNER; LIN & YU; *etc.*)
- TYPE-II HYBRID FEC/ARQ WITH CODE COMBINING (CHASE *et. al.*; KALLEL & HACCOUN, *etc.*)
- GENERALIZED HYBRID ARQ, USING (mk,k) FEC CODES, $m \geq 2$, (KRISHNA, MORGERA, ODUOL)
- RATE-COMPATIBLE PUNCTURED CONVOLUTIONAL CODES WITH ARQ (HAGENAUER; KALLEL & HACCOUN)

SHEET #5

CSI WORKSHOP · 5/16/89 (4)

ADAPTIVELY EQUALIZED MODEMS FOR HF

- SINGLE-TONE MPSK MODULATION WITH FORWARD-ERROR-CONTROL CODING
- USE CHANNEL EQUALIZATION (*e.g.*, DFE) OR COMBINED CHANNEL & DATA ESTIMATION (DDE)
- EQUALIZER PROVIDES A BUILT-IN QUALITY METRIC (MEAN SQUARE ERROR)
- CAN BE MERGED WITH ADAPTIVE CODING

SHEET #6

CSI WORKSHOP · 5/16/89 (4)

scheme in which only error-detection parity checks are sent when the channel is good, and error-correction parities are sent when the channel is bad. In a refinement called Generalized Hybrid ARQ, each retransmission is used to send new parity blocks. In other words the parities that are sent on the resend are parities that actually change the structure of the code with each arriving new set of parities to improve the distance of the code. Work had been done early on by Metzner, Krishna, Morgera, and Haccoun, and some other people in Canada have been further developing those ideas.

The last item I put on here, it's relatively new work and I find it very interesting. Beginning with the work of Hagenauer looking at rate-compatible punctured convolutional codes that have a lot of interesting properties and are decoded with Viterbi decoders. The idea here is that the design is based on what's called a mother code, which is the lowest rate code that you will use in the channel, and higher rate codes are created by puncturing symbols out of the stream that leaves the encoder. This technique has a lot of interesting characteristics. One of them is you can do the puncturing in a way that the bits that are delivered by the decoder on the receive end are delivered with different levels of reliability. In other words, the post-decoding bit error rate is actually different for different bits leaving the decoder. Some work has been done by Hagenauer and some people at ATT Labs.

Looking at, for example, applying that technique to coded speech transmission with any one of the waveform coders, I think it's well known that different bits in the encoded stream are more or less critical for reconstruction of the speech. True with sub-band decoders, with the LPC family, etc. So the marrying of that characteristic of the source with the rate-compatible punctured codes and their properties is a very interesting one. Hagenauer has done some work in wrapping ARQ around that kind of technique.

Let me draw one more little sketch here. What a number of these refinements of hybrid ARQ are doing can be kind of summarized very briefly. What you're trying to do - and I'll just say this is channel quality to make a crude sketch and code rate here. Each of these designs is based on some code which is ... you can think of it this way ... of a set of codes that ranges between some maximum rate and some minimum rate. Depending upon how many times you're willing to resend, that rate can be some very low number down near zero. You're in some way trying to design a technique that will gracefully match the code rate to the channel quality. But the point - and a lot of these refinements here, this and this, for example, have to do with providing some finer intermediate steps, whereas for example simple code combining would go from rate 1/2 down to rate 1/4 and down to rate 1/6. Some of these other techniques provide for a finer grid, if you will, of code rates. But the point I want to get to is that when you commit yourself to a design like this, you're living with some maximum code rate. One of the aspects of simply trying to use hybrid ARQ to adapt to channel quality is that you can certainly see when the channel is getting poorer because you're detecting errors. When the channel goes into a region of good quality, the decoder doesn't really give you information as to how much better the channel is getting. In other words it doesn't really tell you how much better you could do in terms of throughput if you were to shift to a completely different code design. I'll come back to that point in a moment.

[SHEET #6] We've heard a lot about adaptive equalizers here, but I just want to make a few comments regarding the considerable work that's been done and being done to design adaptively equalized modems for radio channels. Again, a lot of attention to HF. They're all based on single tone, M-ary, PSK modulation, and they all have some form of coding incorporated into them. They use – and I'm making a distinction here – channel equalization, for example decision feedback, or combined channel and data estimation. Let me just say briefly what I mean.

I had a discussion with John Proakis after his talk this morning. [SHEET #7] John included this in the category of decision feedback equalization, but it's really different. It's not really equalization. This is a form that has been implemented that has resulted in a very effective modem design. [SHEET #8] Here's a preamble and, in simple terms, the transmission format is alternating blocks of training data and source data symbols. In the nutshell what's going on is that the blocks of training data are used to estimate the channel using, for example, a classical LMS technique, Levinson recursion for example. Then when the channel impulse response is known, the data block is estimated again by LMS estimation. So what's really being done here is channel estimation and data estimation, and the algorithm flips back and forth between the two.

My point in bringing this up is that the reason that this particular technique was devised was that it was based on the observation that on HF channels, for example, there can be some very abrupt changes in the channel impulse response. An equalizer has a rough inverse correspondence of channel transfer function to the the equalizer response. (We know that it's not an exact inverse relationship but roughly speaking it's a good way to think about it.) A very small change in the channel structure can result in a large change in the equalizer, and the equalizer tracking algorithm has the job of rapidly adjusting to that. This scheme (Data-Directed Estimation) gets around that problem, and it's been demonstrated with simulations that this technique in fact does a better job of tracking rapid changes in the channel structure. I don't think I want to take any more time with that. I'll just leave with the comment that the reason John Proakis didn't call this out as a special technique is that asymptotically when it converges the performance will be the same as a decision feedback equalizer. But the tracking characteristics are different and the reason that I bring that up is that even the best of equalizer designs will once in a while lose track of the channel and result in bursts of errors. I'll come back to that point in a minute.

The last two points I wanted to make here is that these equalizers provide a built in channel quality metric. They provide a mean square error that's used in the tracking algorithm and other information that could be used, for example, to derive a signal-to-noise ratio estimate. So that kind of a built-in monitor of channel quality can then be used as a tool to help with the adaptive coding.

[SHEET #9] I'm trying to indicate here in one simple diagram the notion of using this capability and the signal processing capability in the adaptive equalizer modem to implement several forms of adaptation on the HF link. The signal-to-noise ratio estimate could be used with a threshold test to retrain the equalizer to get out of the really bad fades that send the equalizer off track. They could also be used to change the modulation alphabet. There are a couple of stan-

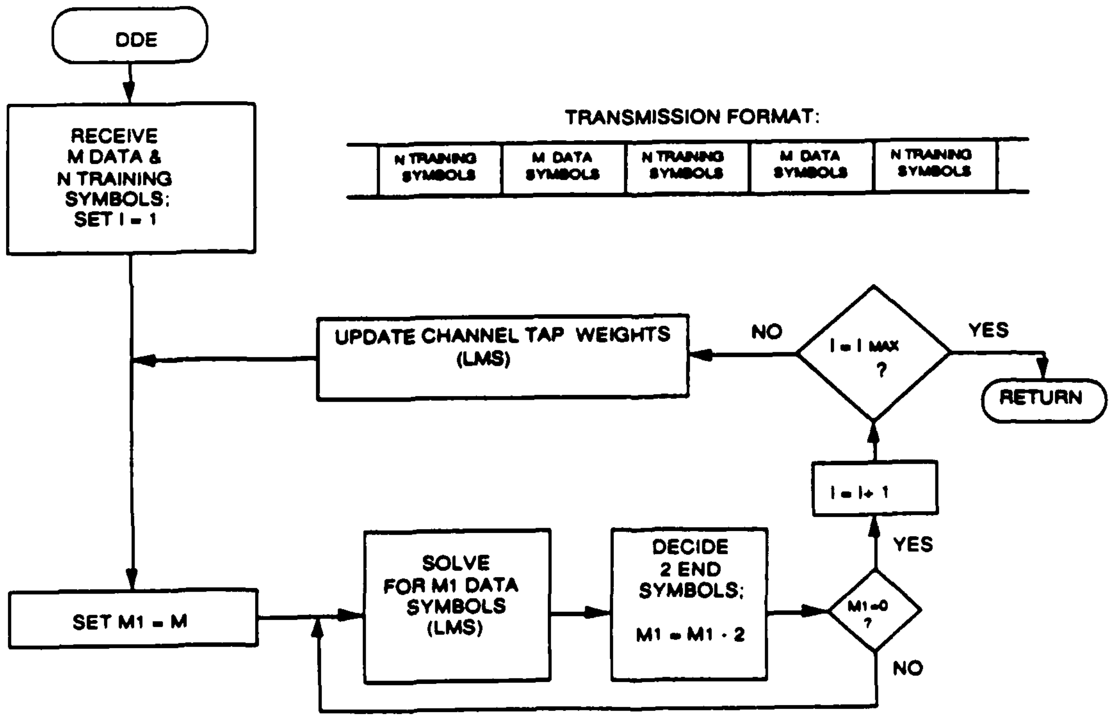
DATA-DIRECTED ESTIMATION (DDE) FOR FADING CHANNELS

- A HIGHLY ROBUST TECHNIQUE FOR FADING MULTIPATH CHANNELS
- KNOWN (TRAINING) DATA BLOCKS ALTERNATED WITH UNKNOWN (SOURCE) DATA BLOCKS
- RECEIVE PROCESSING ALTERNATES BETWEEN CHANNEL ESTIMATION AND DATA ESTIMATION
- ESTIMATION LESS SENSITIVE TO RAPID CHANNEL CHANGES THAN EQUALIZATION
- EXPERIENCE WITH DDE HAS INFLUENCED TWO NEW MODEM STANDARDS (NATO STANAG 4285 & MIL-STD-188-110, REV 2)

SHEET #7

CSI WORKSHOP - 5/16/89 (4)

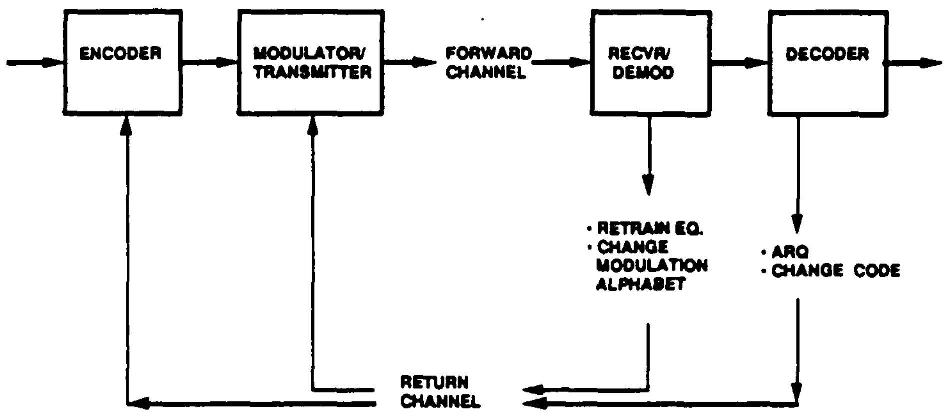
DATA-DIRECTED ESTIMATION



CSI WORKSHOP - 5/16/89 (4)

SHEET #8

MODEL FOR COMBINED ADAPTIVE EQUALIZATION AND ADAPTIVE CODING



SHEET #9

MODES OF ADAPTATION

CONDITION	ADAPTATION
• SHORT-TERM FADING AND IMPULSE NOISE	• HYBRID ARQ (ADAPTIVE)
• SEVERE FREQUENCY-SELECTIVE FADING (EQUALIZER DROPOUT)	• RETRAIN EQUALIZER ON REQUEST (WITH GO-BACK-N)
• LONG-TERM FADING OR CHANNEL DEGRADATION	• CHANGE DATA RATE (MODULATION ALPHABET), AND/OR CHANGE CODE

SHEET #10

SUMMARY

NEW DEVELOPMENTS IN ADAPTIVE CODING (IMPROVED FORMS OF HYBRID FEC/ARQ) AND ADAPTIVE EQUALIZATION FOR FADING/MULTIPATH CHANNELS CAN BE USED TOGETHER TO PROVIDE HIGHLY EFFICIENT USE OF HF RADIO NETWORKS.

SHEET #11

BIBLIOGRAPHY

FADING AND MULTIPATH CHANNELS

- *Special Issue on Fading and Multipath Channel Communications, IEEE* Journal on Selected Areas in Communications, Vol. SAC-5, No. 2, February 1987. (See following two references.)

- S. Stein, "Fading channel issues in system engineering," op. cit., pp. 68-89.

- K. Brayer, "HF data transmission: lessons from the past, directions for the future," op. cit., pp. 90-101.

FEC/ARQ AND ADAPTIVE CODING TECHNIQUES

- D. M. Mandelbaum, "Adaptive-Feedback coding scheme using incremental redundancy," IEEE Transactions on Information Theory, Vol. IT-20, No. 3, pp. 388-389, May 1974

- P. Sindhu, " Retransmission error control with memory," IEEE Transactions on Communications, Vol. COM-25, No. 5, pp. 473-479, May 1977.

- J. J. Metzner, "Improvements in block retransmission schemes," IEEE Transactions on Communications, Vol. COM-27, No. 2, pp. 524-532, February 1979.

- S. Lin and P. S. Yu, "A hybrid ARQ scheme with parity retransmission for error control of satellite channels," IEEE Transactions on Communications, Vol. COM-30, No. 7, pp. 1701-1719, July 1982.

- Y.-M. Wang and S. Lin, "A modified selective-repeat type-II hybrid ARQ system and its performance analysis," IEEE Transactions on Communications, Vol. COM-31, No. 5, pp. 593-608, May 1983.

- S. Lin and D. J. Costello, Jr., Error Control Coding: Fundamentals and Applications, Prentice-Hall, 1983 (Chapter 15)

- S. Lin , D. J. Costello, Jr., and M. J. Miller, "Automatic-repeat-request error-control schemes," IEEE Communications Magazine, Vol. 22, No. 12, pp. 5-17, December 1984.

- G. Benelli, " An ARQ scheme with memory and soft error detectors,"IEEE Transactions on Communications, Vol. COM-33, No. 3, pp. 285-288, March 1985.

- D. Chase, "Code combining - a maximum-likelihood decoding approach for combining an arbitrary number of noisy packets," IEEE Transactions on Communications, Vol. COM-33, No. 5, pp. 385-393, May 1985.

- D. Chase, P. D. Mullers and J. K. Wolf, "Application of code combining to a selective-repeat ARQ Link," presented at 1985 IEEE Military Communications Conference (MILCOM-'85), Boston, MA, October 20-23, 1985.

- C. Lau and C. Leung, "Performance analysis of a memory ARQ scheme with soft decision detectors," IEEE Transactions on Communications, Vol. COM-34, No. 8, pp. 827-832, August 1986.

- J. Hagenauer, " Hybrid ARQ/FEC protocols on fading channels using rate compatible punctured convolutional codes," presented at 1987 IEEE International Conference on Communications, Seattle, WA, June 7-10, 1987.

- H. Krishna and S. D. Morgera, "A new error control scheme for hybrid ARQ systems," IEEE Transactions on Communications, Vol. COM-35, No. 10, pp. 981-990, October 1987.

- S. Kallel and D. Haccoun, "Sequential decoding with ARQ and code combining: a robust hybrid FEC/ARQ system,"IEEE Transactions on Communications, Vol. COM-36, No. 7, pp. 773-780, July 1988.

- S. Kallel and D. Haccoun, "Generalized hybrid type-II ARQ scheme using punctured convolutional coding," in review, 1989.

- J. Hagenauer, N. Seshadri and C-E. Sundberg, "The performance of rate-compatible punctured convolutional codes for digital mobile radio," in review, 1989.

DECISION-FEEDBACK EQUALIZATION FOR FADING CHANNELS

- P. Monsen, "Feedback equalization for fading dispersive channels," IEEE Transactions on Information Theory, Vol. IT-17. No.1, pp.56-64, January 1971.

- D.R. Kern, P. Monsen and C. J. Grzenda, "Megabit digital troposcatter subsystem," presented at 1975 National Telecommunications Conference, New Orleans, LA, December 1975.

- F. M. Hsu, "Square root Kalman filtering for high-speed data received over fading dispersive HF channels," IEEE Transactions on Information Theory, Vol. IT-28, No. 5, pp.753-763, September 1982.

- P. H. Anderson, F. M. Hsu and M. N. Sandler, "A new adaptive modem for long haul HF digital communications at data rates greater than 1bps/Hz," presented at 1982 IEEE Military Communications Conference (MILCOM-'82), Boston, MA, October 17-20, 1982.

- E. Eleftheriou and D. D. Falconer, "Adaptive equalization techniques for HF channels," IEEE Journal on Selected Areas in Communications, Vol. SAC-5, No. 2, pp. 238-247, February 1987

DATA-DIRECTED ESTIMATION FOR FADING CHANNELS

- D. D. McRae, R. S. LeFever and J. N. York, "HF modem feasibility demonstration," RADC-TR-2-48, Final Technical Report, March-October 1981.

- P. Monsen and L. C. Perry, "HF modem test and evaluation," RADC-TR-3-19, Final Technical Report, January 1983.

- G. J. Luhowy and F. A. Perkins, "Advances in HF modem technology," presented at Electronics for National Security Expo, Brussels, Belgium, September 1983.

- F. M. Hsu, "Data directed estimation techniques for single-tone HF modems", presented at 1985 IEEE Military Communications Conference (MILCOM-'85), Boston, MA, October 20-23, 1985.

- J. M. Perl, A. Shpigel, and A. Reichman, "Adaptive receiver for digital communication over HF channels", IEEE Journal on Selected Areas in Communications, Vol. SAC-5, No. 2, pp. 304-308, February 1987.

REAL-TIME EVALUATION OF RADIO CHANNELS

- B. D. McLarnon, "Real-time channel evaluation in an automatic HF radiotelephone system," presented at IEE Second International Conference on HF Communication Systems and Techniques, February 1982.

- J. M. Perl, "Channel coding in a self-optimizing HF modem," presented at International Zurich Seminar on Digital Communications, March 6-8, 1984.

- J. Dawson and M. Darnell, "An HF system design with embedded channel evaluation," presented at IEE Third International Conference on HF Communication Systems and Techniques, London, February 26-28, 1985.

- J. M. Perl and D. Kagan, "Real-time HF channel parameter estimation," IEEE Transactions on Communications, Vol. COM-34, No. 1, pp. 54-58, January 1986.

STANDARDS AND SPECIFICATIONS FOR HF MODEMS

- NATO Standards Agreement (STANAG) 4285, "Characteristics of 1200/2400/3600 bits per second single tone modulators/demodulators for HF radio links," 1 July 1987.

- MIL-STD-188-110, "Equipment technical design standard for common long haul/tactical data modems," Change Notice 2, 4 November 1988.

dards being written. There is a version of MIL-STD-188-110 that will call out a single tone modem with a number of parameter selections. (See Bibliography.) There is also a NATO standard working its way through the approval process that will call out a transmission format for adaptive equalization. I know the Army has a program in which they have funded some developments and implementations of adaptive equalizer modems. Those will all be built with a number of selectable modulation alphabets. So changing the modulation alphabet is certainly one of the things that can be done. In addition to doing simply hybrid ARQ, I'm suggesting the possibility of actually using an adaptive algorithm to change the code selection as well.

The other point that I would want to make is one that has been made by a couple of people already, and that is that even though there's a coder built into the modem, that is not going to help you keep track of the channel. Because, for the reasons that have been discussed, you're not going to take the output symbols from the decoder which may, in fact, also be coming out of an interleaving buffer which further increases the delay. You're not going to try to bring those back and feed them into the equalization process. So even though you may have a very effective coder that has been selected to operate with a given equalizer mode, that coder really doesn't help you keep track of the channel.

I'll finish up with this one. [SHEET #10] What I'm suggesting here is looking at multiple modes of adaptation in which the kinds of things we would ordinarily do such as hybrid ARQ might be done to deal with the short-term channel characteristics (fading, impulse noise, short-term interference, some forms of jamming in frequency hop systems). The more severe problems in the channel, the deep and prolonged frequency selective fades that might cause the equalizer to drop out, can be dealt with by retraining. The very long term characteristics, when the channel at that particular frequency begins to degrade, perhaps the use of that channel can be prolonged by changing the data rate through selection of modulation alphabet and/or changing the code.

I had a few more slides but because I've run out of time I will just look for my summary slide. [SHEET #11] So I'm suggesting this is an interesting and perhaps productive area for new work: looking at ways of merging the processing capabilities and the algorithmic capabilities that are going to exist in these new adaptive equalizer radio modems, and work this together with adaptive coding to do a better job of making use of HF channels and eventually HF networks. Thank you.

ROBERT PEILE: *Adaptive Channel Modeling Using Hidden Markov Chains*[1]

Abstract

We consider the use of hidden markov models and the forward-backward estimation algorithm for real-time channel modeling and, hence, for automated adaptive forward error correction coding. Extensive computer simulation results are presented and interpreted. Design issues for adaptive protocols are exposed and discussed. The primary focus is on digital errors typical of poor quality HF skywave communication and use is made of work on markov modeling of these channels by previous workers [4-6]. Modifications for less disturbed channels are discussed.

1. Introduction and Discussion

We consider the use of hidden markov modeling and the forward-backward algorithm [7]

[1]This work was supported in part by the U.S. Army Research Office under Contract No. DAAL03-88-K-0168.

for adaptive channel modeling and, hence, for adaptive coding. Before presenting the results and details of the models, we consider the motivation for adaptive coding and the present state of coding technology.

Presently, a system designer is faced with a large variety of commercially available coding techniques. In regard to the choice, the standard advice is to "know thy channel" and pick a code accordingly. This ignores the practical point that few system designers know their channel perfectly or that there might not be a single system-wide stationary channel even approximately knowable! Although this was not the case for the traditional breakthrough areas of coding, e.g. commercial satellite links, the problem of apt code selection is becoming acute in both commercial and military communications. In the commercial scenario, data communications are being extended to mobile communications, e.g. the Automated Train Control System (ATCS) [50], data over FM radio and cellular radio. There are obvious problems when the channel conditions are also variable and the noise levels prevent ARQ techniques alone from being adequate. In the military scenario, the spread of data communications to tactical as well as strategic communications can change the frequency range, the mode of propagation and increase the amount of deliberate or accidental interference. Given a diversity of channel conditions and the availability of multicode implementations, the logical response is to select a family of codes capable of meeting the majority of cases in which communication is required. This raises a control issue: how is the most suitable member of the family selected? Even for specialists, manual selection can be difficult on "real" channels. From the point of view of efficiency and ease of use, automation is desirable and, in some arenas, essential. For example, conversation with military instructors indicates that increased training time for non-specialist military operators is expensive and impractical to the point of making operationally complex tactical equipment unacceptable. Faced with growing technical sophistication and with a need to reduce operational complexity, the need for fully automated adaptive systems arises.

There are many different definitions and possibilities for adaptive coding systems. In this article, an adaptive coding scheme has the form depicted in Figure 1. There is a single duplex link between two communicating agencies A and B. (In fact, adaptive coding can be considered on the network level and the half-duplex level. Although this is not the case considered in this article, the results are relevant for these other scenarios.) The traffic is encoded by A, sent across the link, possibly corrupted by noise and decoded at B. To correct errors, it is necessary to locate errors and such location data is valuable information on the link performance. This location data, which may be viewed as self-derived side-information, is fed to a database of statistics. The database is processed in order to get a detailed verdict or update on the statistical mechanism producing the noise. Upon the basis of the noise classification, A and B negotiate and agree upon the most appropriate code for the channel in the A to B direction. An identical process negotiates the most appropriate code in the reverse direction.

It is important to note that adaptive coding works in a different fashion on different channels. Some channels are noisy but statistically stationary, i.e. the conditions are not time-varying. For such channels adaptive coding amounts to an automatic configuration upon activation. Other channels are

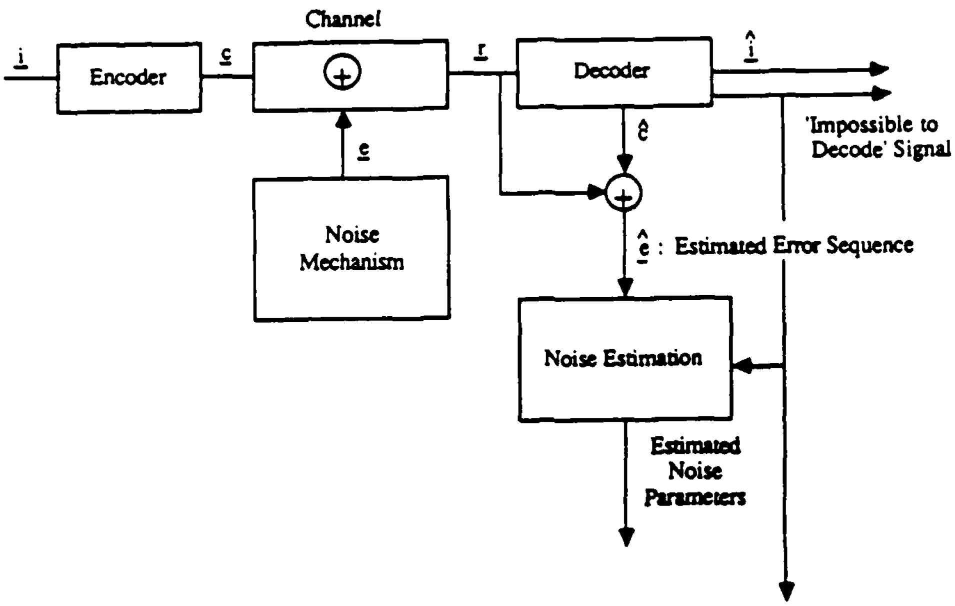
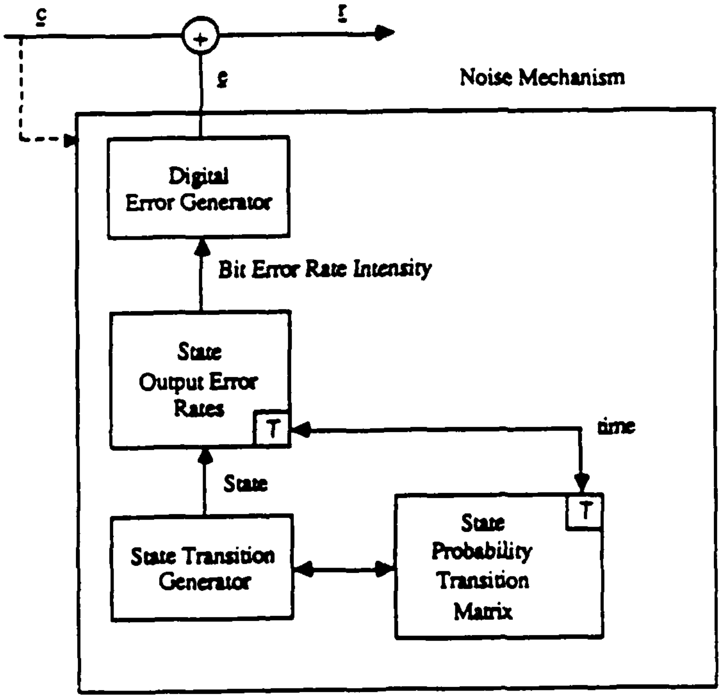


Figure 2: Additive Error Channel Definition



$\boxed{T}$ denotes time-varying parameters.

Note: The time variations might reflect the statistics of the channel data.

Figure 3: Noise Mechanism Model

not stationary and the quality worsens and improves with time, e.g. the notorious HF skywave propagation media. In this environment, adaptive coding is constantly checking and changing the strength of the code. In such extreme circumstances, a fixed code will not be even approximately suitable; if the code were chosen to match one of the many "worst cases," it will be extravagantly redundant for the benevolent periods of channel quality. Conversely, if it is suited for the good periods, it will collapse during the poor periods. Neither extreme is acceptable and the code needs to be changed in reaction to the conditions. If the decoders are working, they have all the information necessary to achieve this adaptation.

Besides desirability, the other motivation is technical feasibility. There are several in-depth articles that chronicle the acceptance and proliferation of error correction techniques and the increasing strength of implementable codes working at practical data rates [50-52]. To give some examples of readily available forward error correcting technology:

1. Qualcomm, Inc., offers a 17 Mbps VLSI convolutional decoder and encoder that has four different code options. Competitors are working on similar 20 Mbps devices.

2. Cyclotomics, Inc., make a unit called the BCM2000 that offers a large array of user selectable options as to Reed-Solomon (RS) code, interleaving options and concatenated options. Although the performance difference between many of the codes is not large, about 100 different RS codes are available for selection on the BCM2000.

3. The growth of memory size and processor power makes the implementation of short binary block codes extremely simple.

4. For harsh channels, *concatenated* codes are well-suited. Concatenated code places two codes in series. The code nearest the channel is known as the *inner* code and is either a short binary code or a convolutional code. By comments 1) and 2) above, there are many choices available for the inner code. The other (*outer*) code is, for sound technical reasons, nearly always an RS code. By comment 2) above, there is a very large choice for this outer code. Consequently, the range of concatenated codes available for selection, even using present day equipment, is large.

Returning to Figure 1, it is immediate that the success of an adaptive coding scheme will rely upon the efficiency of the method used to monitor the error data and classifying the noise mechanism. The method reported upon is that of *hidden markov modeling.* Section 2 describes the method of hidden markov modeling, presents a rational for the selection of this technique and describes the application of the powerful forward-backward algorithm for finding an estimate of the stationary hidden markov model and best approximates a finite sequence of channel errors. Section 3 describes the codes and channel models selected for simulation and Section 4 presents results. Section 5 discusses the results and system issues. Section 6 discusses areas for future work. Section 7 presents conclusions.

2. Hidden Markov Modeling

2.1. Description of the channel and simulation model

Figure 2 shows the additive error channel that was investigated. A digital data

source $\underline{i}$ is encoded to give a digital stream $\underline{c}$. This stream is transmitted over a channel. A noise mechanism generates a digital sequence $\underline{e}$ which is added (exclusive-or addition) to $\underline{c}$ to give a digital stream $\underline{r}$. A decoder receives the stream $\underline{r}$ and processes it. One output of the decoder is a sequence $\hat{\underline{i}}$, the decoder's estimate of the transmitted data $\underline{i}$. $\hat{\underline{i}}$ is passed to the data sink. Another output is an "impossible-to-decode" flag that allows the decoder to alert external circuitry of a detected glitch in the decoding process. The third output is an estimate $\hat{\underline{c}}$ of the transmitted coded sequence $\underline{c}$. This can be added to $\underline{r}$ to give an estimate $\hat{\underline{e}}$ of the error sequence introduced by the noise mechanism. (In fact, several decoding algorithms give $\hat{\underline{e}}$ directly. Others might need a re-encoder to obtain $\hat{\underline{c}}$ or $\hat{\underline{e}}$.) The strategy of the proposed adaptive coding scheme is to process $\hat{\underline{e}}$ to obtain an estimate of the noise mechanism parameters. The estimate is used to predict the severity of forthcoming noise and to select a code accordingly, (see Figure 1b).

Figure 3 shows the hypothesized model of the noise mechanism. The channel is assumed to have a finite and fixed number of states. At each bit interval the channel selects a state according to the probabilities in a state transition matrix. The selected state is an address to a memory of bit error rate probabilities. The output BER probability is an input to a digital random error generator which generates either an error (1 in $\underline{e}$) or a non-error (0 in $\underline{e}$) with the choice governed by the value of the BER. Note:

1. The state output error rates and the state transition matrix can be time-varying and/or dependent upon the statistics of the channel data $\underline{c}$.

2. The digital error generator is memoryless; the channel's memory is embedded in the markov state sequence.

The simulations made the assumption that the decoders were supplying an accurate estimate of the errors, i.e. that $\hat{\underline{e}} = \underline{e}$. With this assumption, which is discussed below, the simulations were of an equivalent additive error channel shown in Figure 4. The noise generator of Figure 3 generates a sequence $\underline{e}$ which is passed (via the decoder) to a process which attempts to estimate the noise parameters. These estimates are passed to a prediction mechanism that attempts to select the best possible code under the assumption that the noise parameters will be static during the transmission of a codeword. As a calibration exercise, the true noise parameters are passed to an identical noise prediction device that selects a code under the same assumption. (This latter prediction is referred to as the "Genie" prediction.)

The above channel is a markov chain λ for which there are N channel states. At time t, the channel state is a random variable Q^t with a sample space equal to space of possible channel states. Transitions between Q^t and Q^{t+1} depend upon an N by N transition matrix which has the i-th, j-th entry equal to the probability of going from the i-th state to the j-th state, $a^t_{i,j}$, say, at that time. We refer to this matrix as A^t. In many markov chains we can directly observe the state of Q^t at time t; in a hidden markov chain this is not the case. A hidden markov chain is a markov chain in which the data that can be observed is a *probabilistic* function of the underlying state sequence. The observable data can only indicate the relative likelihoods of the possible state sequences. In the general model, suppose that we observe a variable $O = O^t$ that can be in one of M states. The probability of $O^t = k$ given that $Q^t = j$ is

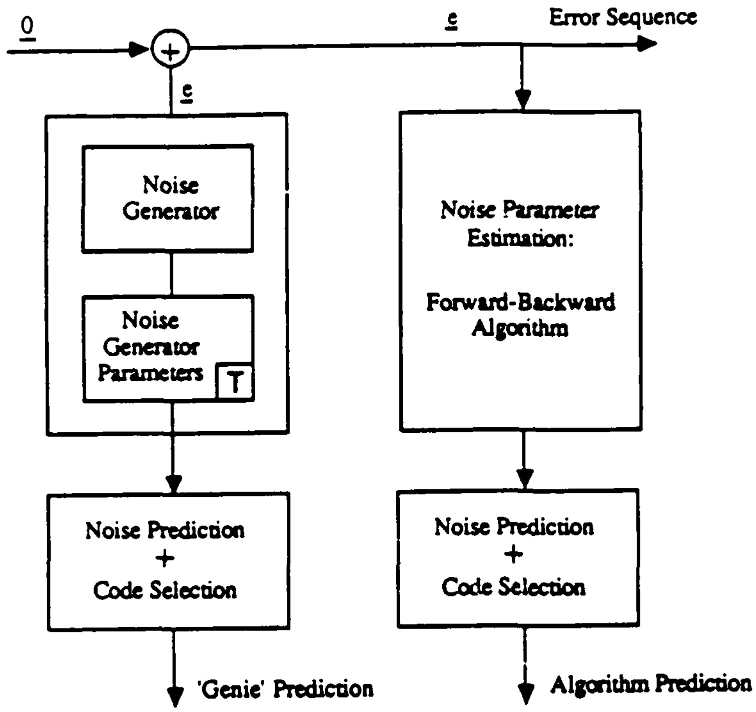
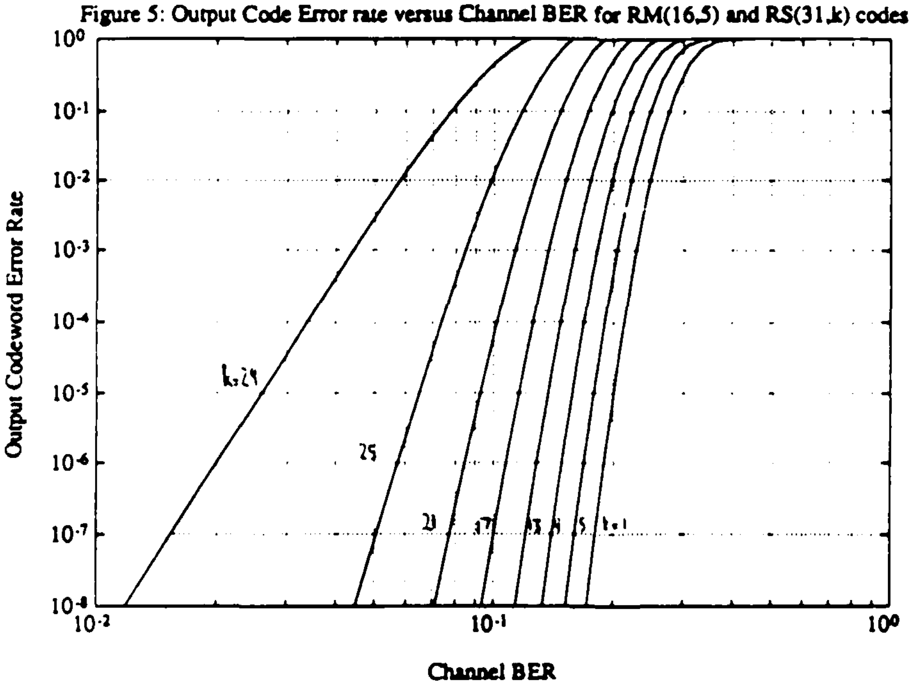


Figure 4: Simulation Model

Figure 5: Output Code Error rate versus Channel BER for RM(16,5) and RS(31,k) codes



equal to $b_j^t(k)$, where $k : 1 \leq k \leq M$ and $j : 1 \leq j \leq N$. We put the matrix of $b_j^t(k)$ equal to B^t.

In this channel model $M = 2$ and O^t equals 0 or 1, i.e. $\underline{e}$ is a sample sequence selected from O. Given that Q^t is in state j, $b_j^t(0)$ is the probability of a non-error and $b_j^t(j)$ is the probability of an error.

Note that, for a stationary hidden markov chain, the 't' superscript may be removed from $a_{i,j}^t$, $b_j^t(k)$, A^t and B^t. A central problem with stationary hidden markov chains is to start with an observed sequence $o_1, o_2, o_3, \ldots, o_T$ of samples or outcomes from O and try to estimate:

1. the most likely state sequence of Q underlying the process;
2. the most likely values for the $a_{i,j}$;
3. the most likely values for the $b_j(k)$;
4. the most likely value for π, the probability of being in each state at time 1 (the initial state probability vector).

The forward-backward algorithm provides an iterative solution to this problem for stationary hidden markov chains, as discussed in Section 2.3. In the channel model, the algorithm supplies an estimate of the noise mechanism from observing $\hat{\underline{e}}$. This estimate's accuracy depends upon the assumptions that the noise mechanism is stationary across the samples and that $\hat{\underline{e}}$ equals $\underline{e}$.

This raises several questions. Firstly, it is not clear how accurate the algorithm's estimates will be if the parameters are time-varying across the entire block of observed data. It is also not clear if the algorithm can function given bad initial estimates caused by violent changes in the noise mechanism. Addressing these issues was the major aim of the simulations. At another level, the assumption of $\hat{\underline{e}} = \underline{e}$ deserves comment. If the assumption is false, the decoder is not coping with the prevailing noise conditions. This, for most decoders, will be very highly correlated with activity on the "Impossible-to-Decode" signal. This signal is seen as an interrupt to the control protocols that prompts the logic to strengthen the coding regardless of the output of the noise estimation process. In other words, if the assumption is false, the functionality of the algorithm is largely irrelevant. (However, the ability of the algorithm to regain stability after an inaccurate estimate is important.)

Finally, the entire model rests upon the suitability of a hidden markov model for errors on a communications channel. A justification for using hidden markov modeling within the adaptive coding context is given in Section 2.2.

2.2. Hidden markov channels and communications channel

Markov chains have been frequently used to model time dispersive communications channels, e.g. HF, tropospheric scatter and microwave. In particular, there are important articles by (in chronological order) Gilbert [1-2], Elliot [3], Fritchman [4], Adoul *et al.* [5] and McManamon [6]. The last three have proved particularly important to this work as they provided data on transition matrices observed on HF channels; these proved useful inputs to the simulations. These articles also contain data on the goodness of fit between observed error data and the best markov model.

Two conceptual differences separates the work in [4] and [6] from this work. Firstly, these workers set up markov chains rather than hidden markov chains. The channel was considered to have N states of which some

fraction always of which some fraction always causes an error and the others always coincide with a non-error. The possibility of being in an unfavorable state and not getting an error does not occur. The hidden markov chain includes this formulation as a special case but takes a more general perspective. The channel is seen as a probabilistic case of errors rather than identical to the appearance of errors. For example, the most prevalent cause of poor performance on HF media is multipath, i.e. the existence of multiple paths with multiple attenuations between transmitter and receiver. One immediate possibility is to model each path combination as a state in a markov chain. With this formulation, the state sequence corresponds to a succession of path combinations. Each path combination will have a different received Signal-to-Noise ratio and, hence, a different Bit Error Rate after demodulation. In fact, this point is academic to the results and performance of the algorithms; the model finds the best markov model fit regardless of physical interpretation. If it was desired to find the best model with state error probabilities of zero or one, this can be achieved by setting an initial estimate for the probabilities as zero or one. Conversely, to retain the full power of the model, this should be avoided.

In summary, the multipath mechanism of HF and other scintillation channels present conditions that it is natural to consider as a time-varying hidden markov model. The time-varying nature of these channels suggest that short-term variations are caused by the markov chain structure and long-term effects by time variations in the markov model. Of course, such a model is only of use if there are implementable algorithms for extracting the best hidden markov model from the observed error statistics. This is the subject of the next section.

2.3. The Forward-Backward Algorithm

An excellent description of the forward-backward algorithm, and its applications can be found in Rabiner and Juang [7]. The algorithm has been applied to problems in speech analysis and vocoding, textural based image recognition and segmentation, seismic data processing and cryptanalysis [9-46]. Given the general nature of hidden markov models it is to be expected that the algorithm will find many more applications. There are strong similarities between the algorithm and that of the Viterbi/Bellman algorithm [8]. This is discussed in [7]. The algorithm can be applied to doubly stochastic processes where the random variables are discrete or continuous. Within the context of this work, we have applied the algorithm to binary data. However, a more sophisticated system might wish to use codes that process soft-decision noise data, e.g. continuous valued random variables or random variables with 8 or 16 output values. The algorithm is capable of such extensions.

The Algorithm (as used in this work) addresses the following problem:

Problem Statement

Given a finite data sequence $\underline{e}$ of '0' and '1's, and given the number of states N, find the parameters of the stationary hidden markov chain with N states and initial state probability vector that maximizes the probability of observing the given data.

Box 1 presents the algorithms. [7] explains their derivation. The algorithm iteratively finds markov models that approximate the observed data. It is possible to show [45] that each new model will lead to a higher probability of observing the given sequence than the previous. More precisely, it was shown in

Box 1: Forward-Backward Algorithm

(A) Input:

$A = (a_{i,j})\ 1 \le i,j \le N$

$B = (b_j(k))\ 1 \le j \le N,\ 1 \le k \le M$

$\pi = (\pi_i),\ 1 \le i \le N$, initial state distribution vector

$0 = (0_1, \ldots, 0_T)$ observed sequence

(B) Iterative Step:

(1) Forwards Step:

1.1 Put $\alpha_1(i) = \pi_i b_i(0_1),\ 1 \le i \le N$

1.2 For $t = 1,2, \ldots, T-1,\ 1 \le j \le N$

$$\alpha_{t+1}(j) = \left[\sum_{i=1}^{N} \alpha_t(i) a_{i,j}\right] b_j(0_{t+1})$$

1.3 Put $P = \sum_{i=1}^{N} \alpha_T(i)$

(2) Backwards Step:

2.1 Put $\beta_T(i) = 1,\ 1 \le i \le N$

2.2 For $t = T-1, T-2, \ldots, 1,\ 1 \le i \le N$

$$\beta_t(i) = \sum_{j=1}^{N} a_{i,j} b_j(0_{t+1}(j))$$

(3) Intermediary Parameters:

3.1 $\gamma_t(i) = \dfrac{\alpha_t(i) \cdot B_t(i)}{P},\ 1 \le t \le T,\ 1 \le i \le N$

3.2 $\xi_t(i,j) = \dfrac{\alpha_t(i) a_{i,j} b_j(0_{t+1}) \beta_{t+1}(j)}{P},\ 1 \le t \le T,\ 1 \le i,j \le N$

(4) Re-estimation Formulas:

4.1 $\bar{\pi}_i = \gamma_1(i),\ 1 \le i \le N$

4.2 $$\bar{a}_{i,j} = \frac{\left[\sum_{t=1}^{T-1} \xi_t(i,j)\right]}{\left[\sum_{t=1}^{T-1} \gamma_t(i)\right]}$$

4.3 $$\bar{b}_j(k) = \frac{\left[\sum_{t=1, 0_t=k}^{T} \gamma_t(j)\right]}{\left[\sum_{t=1}^{T} \gamma_t(j)\right]}$$

(C) Repetition:

Replace A, B, π by $\bar{A}, \bar{B}$ and $\bar{\pi}$ and repeat Step B.

the Appendix A of [45] that:

$$Pr(\underline{e}/\lambda*) \geq Pr(\underline{e}/\lambda),$$

where λ is the old model estimate and $\lambda*$ is the new model estimate, ie. the re-estimation guarantees a better model.

This indicates a highly desirable general property of the algorithm but the performance still carries caveats. If the initial estimate is in a "saddle-point" region, the rate of climb of probability versus iteration can be impractically slow. This can be compensated; one modification that has been suggested involves monitoring the rate of climb. If the probability and rate of change are both unacceptably small, the process is forced to "shoot," i.e. change by a large in order to avoid the local trap [48]. In fact, in the experiments involving channel models, a very few number of iterations were necessary for convergence. It is also possible [45], in the general case, for a bad initial estimate to mislead the estimation process for some period of time, i.e. many iterations are necessary. Again, in the experiments involving channel models this seems less a factor than for the general case.

The output of the algorithm, when there are disallowed transitions, may not be a valid state sequence. The solution determines the most likely state at every instance. An alternative solution would be to use the Viterbi algorithm to find the most likely path or state sequence [8]. To clarify the difference it is interesting to consider the advantages and disadvantages that would result from applying the forward-backward algorithm to a convolutional decoder instead of the Viterbi Algorithm. The Viterbi Algorithm would produce the most likely state sequence. The forward-backward algorithm produces the sequence of states that gives the lowest difference to the observed data. In other words, whereas the Viterbi Algorithm is, in some sense, a global solution, the above produces a series of local solutions.

Another practical factor relates to scaling. The estimation quantities rapidly diminish in size and, in a finite computer, some corrective action must be taken. This is discussed in [45].

3. Channel Model Experiments

3.1. General description

The efficiency of the described algorithm was examined by computer simulations. The major elements of a simulation included:

1. Input of the types of code to be adapted;
2. Input of the underlying noise mechanism that it is wished to estimate;
3. Generation of noise sample sequences in accordance to the input noise mechanism;
4. Estimation of a hidden markov model (HMM) that fits the noise sequence using the forward-backward algorithm;
5. Prediction of the best code to use given the estimated HMM;
6. Prediction of the best code to use given that the transmitter is told the true value of the HMM at the start of each codeword, i.e. a "genie"-based prediction;
7. Output of results.

These are addressed below.

Types of codes

The particular class of codes that was emphasized concatenated an $RS(31,k)$ code on 5-bit symbols with that of a (16,5) binary Reed-Muller code. The Reed-Muller codeword is used to protect each symbol of the

RS code. The code is three binary error correcting and four binary error detecting. The input to the RS decoder is mostly correct or flagged as an erasure; symbol errors are in the minority. The $RS(31, k)$ codes are allowed to vary k from 1 to 31; this is the parameter that is the target of the adaptive coding. The code redundancy, as presented in the results, is defined to be $(31 - k)/31$ where k is the selected redundancy. (For the true redundancy, this has to be increased.) The codes are strong at combating both burst and random errors. Figure 5 shows the performance against random errors. Note that error rates of worse than 1 in 10 are correctable with high probability. (This class of codes, with k manually selected, was implemented using DSP chips by one of the authors in 1982 and used successfully with HF serial modems.)

Input parameters of the noise mechanism

In terms of channels, the experiments used cyclostationary hidden markov models with period N as a noise mechanism, where N is a variable. (This is described in 3.2.) The parameters were selected from published literature where possible and supplemented where necessary. All the noise was binary; no experiments with "soft" Gaussian noise or quantized inputs were performed.

Noise generation

The noise generation was based around standard random number generation routines. It is well known that these models are not ideal in terms of the higher order properties of a truly random number generator. However, these properties were not regarded as critical to the experiments. And, if pertinent at all, constituted an additional challenge to the HMM algorithm.

Hidden markov model estimation

The computer model took a codeword of noise and called the HMM estimation algorithm on a once per codeword basis. For the codes discussed above, this constitutes a length of $31.16 = 496$ bits duration. In practice, the noise data would be obtained from the error-correcting decoder and it is logical to call the estimation routine as the data becomes available. However, there is no reason that the HMM algorithm needs to be tied to the codeword size, as discussed in Section 5.

Estimated code selection

Given an HMM noise mechanism across a codeword, code performance can be predicted by driving two further markov chains from the HMM. Firstly, a markov chain is used to compute the distribution of the number of errors that occur across the 16 bits of a Reed-Muller codeword. This gives the probability of a correct, erased and erroneous symbol. A second markov chain then computes the distribution across the 31 symbols of the RS code of the quantity:

2.(Number of symbol errors) + (Number of symbol erasures).

This last distribution leads to the probability of the $RS(31, k)$ codes concatenated with the RM code being able to correct in the presence of the HMM. In the experiments this process was performed using the estimate obtained from the last codeword as an initial estimate of an HMM stationary across the subsequent codeword. With this approximation, the maximum value of k was selected that made the probability of codeword error below a threshold set to 10^{-4} in the experiments.

Genie code prediction

An approach was used whereby the model generated the code that the transmitter would use if perfect a priori information was supplied by a benevolent "genie," see Figure 4. In more detail, a code was selected using

the true value of the HMM at the start of each codeword. This HMM was used to drive the two markov models described above. Using the same 10^{-4} threshold and presuming that the model remained unchanged across the codeword, a code was selected. This code represents the best choice for a transmitter lacking hindsight or sophisticated circuitry predicting the rate of change of markov chains.

Output of results

Results were processed to give a graphical view of how the selected redundancy varied with time. This is the subject of Section 4.

3.2. Channels and noise generation

Noise data was generated using cyclostationary markov chains. The model used the weighted sum of two fixed markov chains where the weights varied sinusoidally with a period equal to that of the cyclostationarity.

More precisely, let $\underline{A}$ and $\underline{B}$ be any two arrays of real numbers of the same dimension. Then, at time t, the array C_t can be constructed where:

$$C_t = a(t) \cdot \underline{A} + [1 - a(t)] \cdot \underline{B},$$

where:

$$\begin{aligned} 0 &\leq a(t) \leq 1 \\ a(t) &= 0.5(1 + \sin(2 \cdot \pi \cdot f \cdot t), \qquad (1) \end{aligned}$$

where f has period N_f, i.e. $f \cdot N_f = 1$. Note that, if the original arrays were non-negative with either row or column sums equal to unity, C_t has the same property. It follows that if $\underline{A}$ and $\underline{B}$ were probability transition matrices, C_t is a probability transition matrix for every possible t and $C_{t+N_f} = C_t$. If $\underline{A}$ and $\underline{B}$ are two possible output BER matrices, i.e. indexed by the channel states and giving the pronability of error and non-error for each state, C_t is also an output BER matrix.

A cyclostationary markov chain with period N_f was constructed by applying this weighted sum combining to all the parameters of the underlying markov model. The input to the model includes:

1. the period N_f;
2. two estimates of the initial state probability vector;
3. two state transition probability matrices; and
4. two vectors giving the BER arising from each state.

Note that making the paired vectors or matrices identical implies that the input is not varied with time, i.e. $\underline{A} = \underline{B}$ implies $C_t = \underline{A}$.

The construction was used in two ways. The principal use was to make the two transition matrices identical and to have the output BER vectors different. The result is to simulate an HF link that is stationary except for a periodic attenuation of the received signal. In other words, the probabilities of changing from state to state (or path to path) are unchanged but the noise afflicting each and every state varies. The other use was to make the transition matrices different. The resulting model is more complicated and was selected to represent a fundamental changing of the multi-path spread with time as well as variable attenuations. The markov models were updated at every other channel bit.

The actual channel parameters were as follows.

Channel 1:

This model used a two state markov chain for the underlying noise mechanism. In this example, the markov chain, state probability vector and output BER vector were not time-varying, giving a stationary noise mechanism.

The transition matrix was set equal to:

$$\begin{bmatrix} 0.8, & 0.2 \\ 0.3, & 0.7 \end{bmatrix}$$

The first state represents a "good" condition and the second represents a "bad" condition in that the probability of a non-error in the two states are 0.99, 0.9 respectively, i.e. the good states gives a BER of 10^{-2} and a bad state gives a BER of 10^{-1}. The simulation run with an initial state probability of 0.5 for each state, i.e. with no preference to which state the experiment started in.

This model tends to produce bursts of data containing a relatively high density of errors separated by longer intervals with a lower density.

Channel 2: Channel 2 was identical to Channel 1 except that the output BERs of the state were made time-varying. At one extreme, the BERs were as in Channel 1, i.e. 10^{-2} and 10^{-1}. At the other (worse) extreme the output BERs of the two states were set as 2.10^{-2} and 2.10^{-1}. In other words, the probability of an error from a state was, at best, 0.01 and 0.1 respectively and, at worst, 0.02 and 0.2 respectively. The actual value at time t was equal to:

$$a(t) \cdot [0.01, 0.1] \; + \; [1 - a(t)] \cdot [0.02, 0.2]$$

where $a(t)$ varies sinusoidally between '0' and '1'. The period of $a(t)$ was set to 4000 bits, i.e. the time-varying chain is identical every 4000 bits. The choice of this value is discussed in the results section.

This model produces a similar effect to model 1 except that the density of errors in both the good and bad sections varies in intensity.

Channel 3:

This model has two states as above. The transition matrix was fixed but the output BER vector was time-varying as for Channel 2. The transition matrix was set equal to:

$$\begin{bmatrix} 0.6, & 0.4 \\ 0.4, & 0.6 \end{bmatrix}$$

The output BER vectors were set to (0.01,0.05) and (0.01,0.1). The weighted sum of these vectors has a period of 4000 bits.

This model was selected to examine the effects of making the lengths of good and bad sequences more comparable.

Channel 4:

This was identical to Channel 2 except for the extremal values of the output BERs. These were set as $(10^{-2}, 10^{-1})$ and $(3.10^{-2}, 3.10^{-1})$. The period was 4000 bits as before.

The model was selected to provide a greater range of conditions than in Channel 2.

Channel 5: Channel 5 used a three state markov chain. The noise mechanism is stationary and uses values taken from [4] which, in turn, were found by analyzing the errors observed on a benevolent HF link. The transition matrix is:

$$\begin{bmatrix} 0.66, & 0.0, & 0.34 \\ 0.0, & 0.9991, & 0.0009 \\ 0.44, & 0.34, & 0.22 \end{bmatrix}$$

In accordance with the Fritchman models, the associated probability of an error was set as either 0 or 1. More precisely the probabilities were set as (0,0,1).

Channel 6:

Channel 6 modified Channel 5 to give time-varying output Bit Error Rates of increased intensity. The selected Output BERs varied between best and worst cases equal to [0.01,0.01,0.5] and [0.05,0.1,0.5]. The period cycle was set equal to 4000 bits.

Channel 7:

Channel 7 has the same Fritchman-derived non-time-varying transition matrix as Channel 6. However, the output BERs were made more dynamic with a more severe worst case. The best case BERs were [0.01,0.01,0.5] and the worst case BERs were [0.9,0.9,0.5]. (The reason for imposing such impossible worst case conditions was to ascertain the speed that the HMM estimation recovered from an extremely pessimistic initial condition.)

Channel 7b:

Channel 7b was identical to Channel 7 except that the period cycle was extended to 40,000 bits, i.e. the conditions were slowly varied from best to worst case.

Channel 8:

Channel 8 took transition matrix values that were extracted from tests on an HF link and published in [6]. The transition matrix was not varied with time and is equal to:

$$\begin{bmatrix} 0.698, & 0.0, & 0.0, & 0.302 \\ 0.0, & 0.9976, & 0.0, & 0.0024 \\ 0.0, & 0.0, & 0.99935, & 0.00065 \\ 0.639, & 0.264, & 0.015, & 0.0082 \end{bmatrix}$$

Unlike [6], the BERs were made time-varying rather than "0" or "1". The best and worst cases were set to [0.05,0.01,0.01,0.5] and [0.1,0.1,0.1,0.5] respectively. The period was 4000 bits.

Channel 8b

This was identical to that of Channel 8 except that the period of time variation was slowed down to 40000 bits.

Channel 9

This channel varied the state transition probability matrices between (essentially) that of Channel 5 and that of Channel 8. One of the matrices was identical to that of Channel 8. The other was equal to:

$$\begin{bmatrix} 0.66, & 0.0, & 0.44, & 0.0 \\ 0.0, & 0.9991, & 0.34, & 0.0 \\ 0.44, & 0.0009, & 0.22, & 0.0 \\ 1.0, & 0.0, & 0.0, & 0.0 \end{bmatrix}$$

The best and worst BER vectors were set to [0.05,0.01,0.01,0.5] and [0.1,0.1,0.1,0.5] respectively. The period was 4000 bits.

3.3. Other parameters and comments

The simulation model used a small number of iterations for the hidden markov model estimation, i.e. three. Experiments were performed with larger values but with no observed return in accuracy. At the *end* of each codeword, two predictions were made as to the code redundancy to be utilized. The first "genie-based" prediction used the actual markov chain prevailing at the start of the next codeword. The second used the HMM estimate computed from the last codeword. This implies that there is a timelag between the genie and the HMM predictions.

For any cyclostationary markov chain it is possible to work out the marginal distribution at a particular time. This allows the average probability of bit error on the channel to be computed at any time. This was computed for the above models and the results are discussed below. However, the channel BER results can be misleading. Being averaged marginal distributions, the BER versus time graphs exhibit no information as to the burstiness of the distributions.

4. Results

Figure 6 shows the average BER versus time graph for Channel 1. The Markov chain is stationary implying that the marginal distribution and average BER are constant. Figure 7 shows the results of the code simulations for Channel 1. The straight line in Figure 7 represents the ideal "genie" prediction. The

other line describes the code redundancy recommended by the HMM estimation process. Clearly the HMM estimation process is centered around the ideal prediction. It is worth noting that the HMM deviations are to be expected, the noise samples over 496 bits can exhibit statistical deviations from the underlying model and make the "best" fit different from the average. Less formally, we can get good and bad stretches of data. Note that a smoothing or averaging process applied to the output would produce a very accurate estimate, albeit at a cost in adaptation to time variations.

Figure 8 shows the average BER versus time for Channel 2. The time variation imposed by the sinusoidally weighted sum is clearly discernable. Figure 9 shows the results of the simulations. There are two lines. The genie prediction can be seen to be an approximation to a periodic sine wave. (The difference between codeword length and fundamental frequency imposes some variation in the genie prediction; on different cycles the predictions are made at slightly different phases.) The HMM estimate is also depicted in Figure 9. The latter is necessarily less smooth than the "genie" estimate but it is clearly tracking the genie estimate. The codeword lag between the two estimates is evident.

Figure 10 shows the averaged BER for model 3. The channel is evidently less severe than Channel 2 as regards BER. However, the transition matrix is structured to produce longer noise bursts. Figure 11 shows the results of simulations. The codes under consideration tend to favor the presence of noise bursts and this is reflected in the reduced redundancy selected by both predictions. Moreover the variation in code redundancy versus time is much reduced. The HMM estimate is clearly tracking the genie prediction accurately and quickly.

Figure 12 shows the average BER versus time for Model 4. This channel is more extreme than the foregoing; the time variation is larger and the worst case conditions more extreme than any of the previous channel models. Figure 13 shows the simulation results. The genie prediction reflects the channel conditions, oscillating between an RS redundancy of about 0.2 to 0.7. The HMM estimate is accurate in that the two lines are closely intertwined. The HMM estimate tends to be pessimistic. For example, worst case redundancies of 0.8 are predicted. This is not necessarily a fault of the algorithm; the model produces bursts. If a particular codeword hits a long burst, the HMM model is bound to produce an estimate more pessimistic than the long-term average. The time lag between the two predictions is clearly visible.

Figure 14 shows the average BER for Channel 5. The channel is not time-varying and produces a channel BER of below 0.003. Simulation results are not exhibited in this case; the majority of codewords had no errors, i.e. the estimation process had no input o.. which to function. In fact, the ideal RS code is no RS code in this case; the RM code can function without assistance. In Section 5 we discuss how the HMM model can be structured to work in low error rate regimes.

Channel 6 consists of Channel 5 with the output Error Rates worsened and time-varying. Figure 15 shows the average BER versus time graph. Figure 16 shows the results of the simulations. The HMM prediction is clearly tracking the genie estimation. The growth of the model from two to three states seems to have little effect on the accuracy or inaccuracy of the model.

Figure 6: BER versus time for model 1

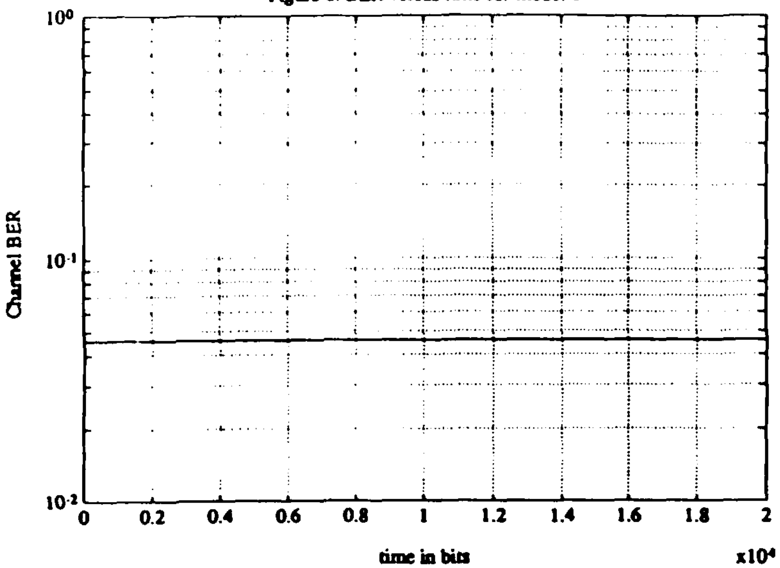


Figure 7: Results for model 1

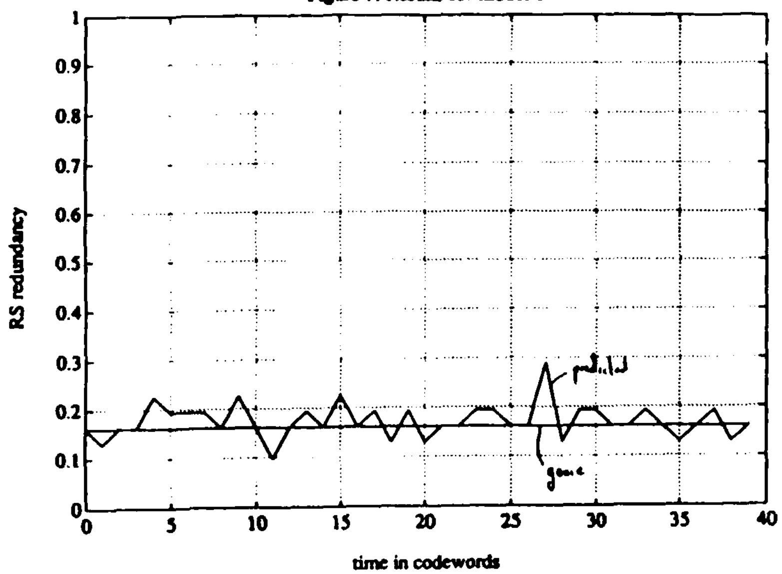


Figure 8: BER versus time for model 2

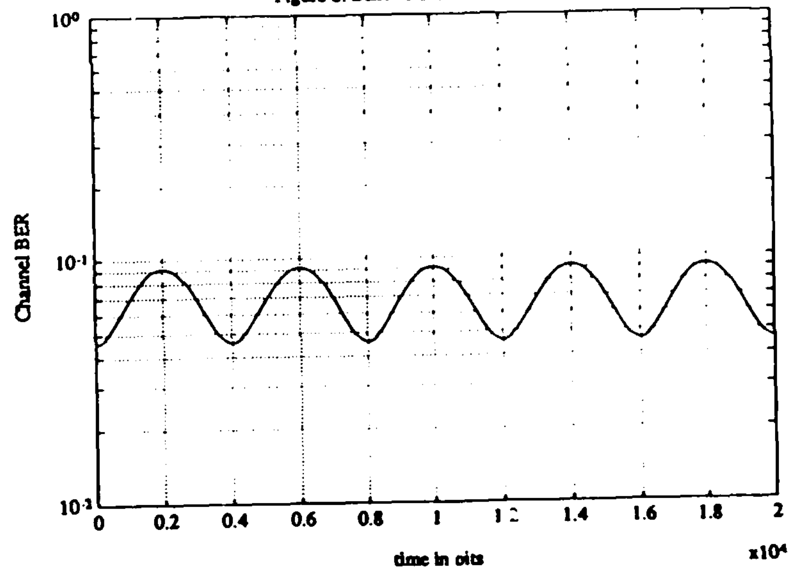


Figure 9: Results for model 2

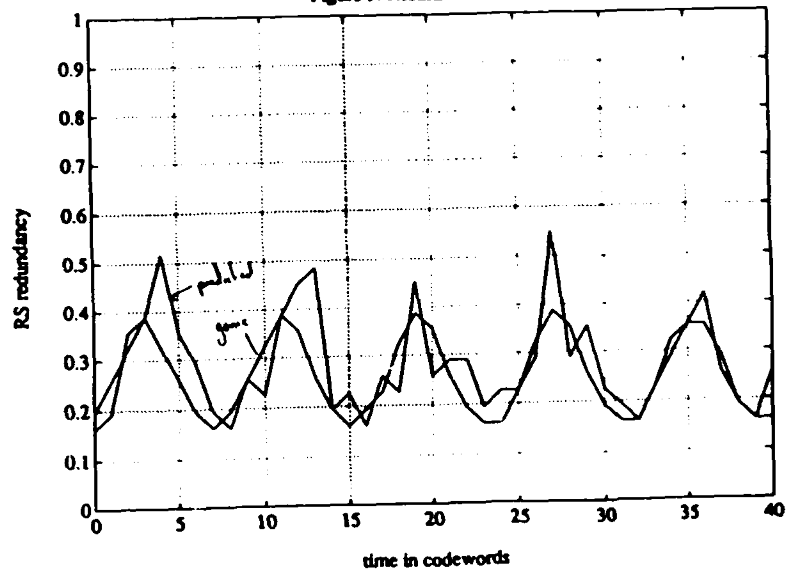


Figure 10: BER versus time for model 3

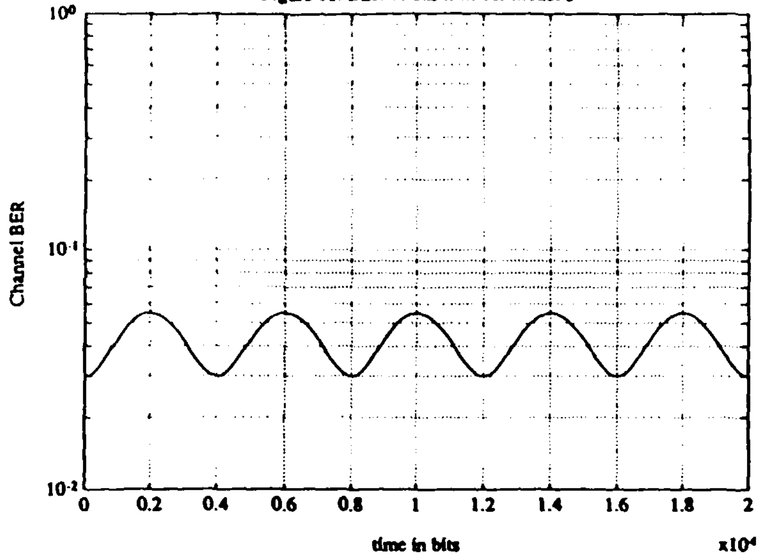


Figure 12: BER versus time for model 4

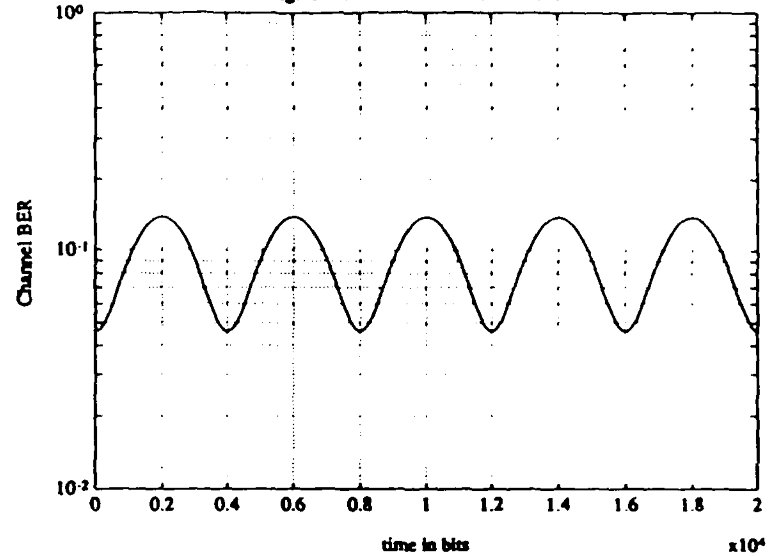


Figure 11: Results for model 3

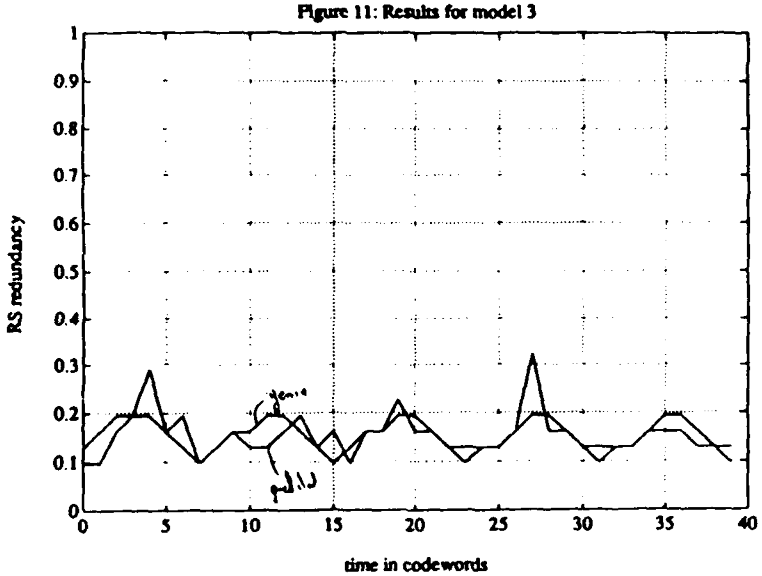


Figure 13: Results for model 4

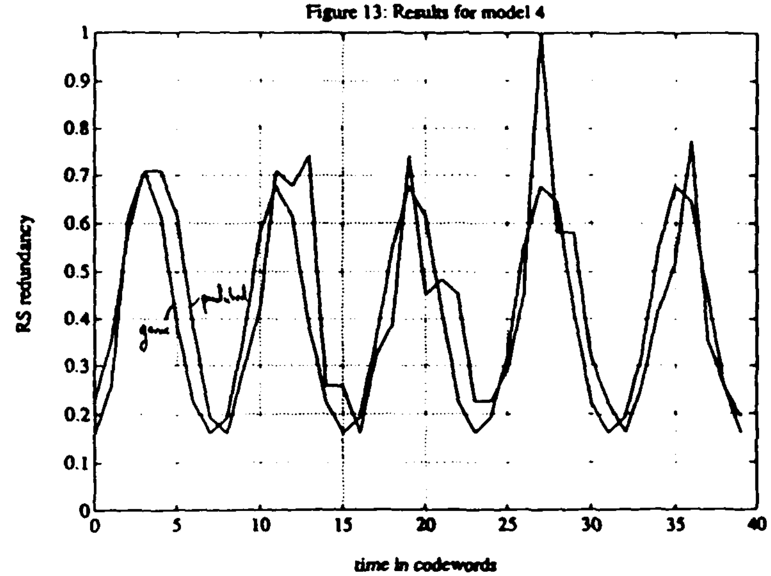


Figure 14: BER versus time for model 5

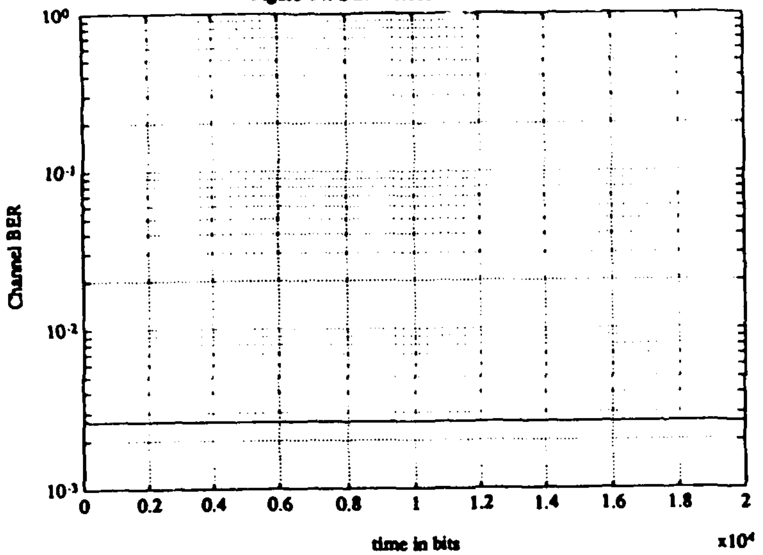


Figure 15: BER versus time for model 6

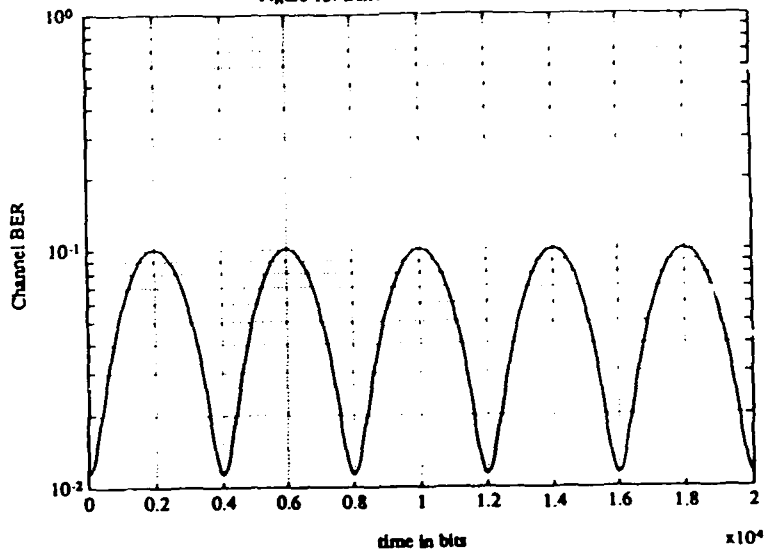
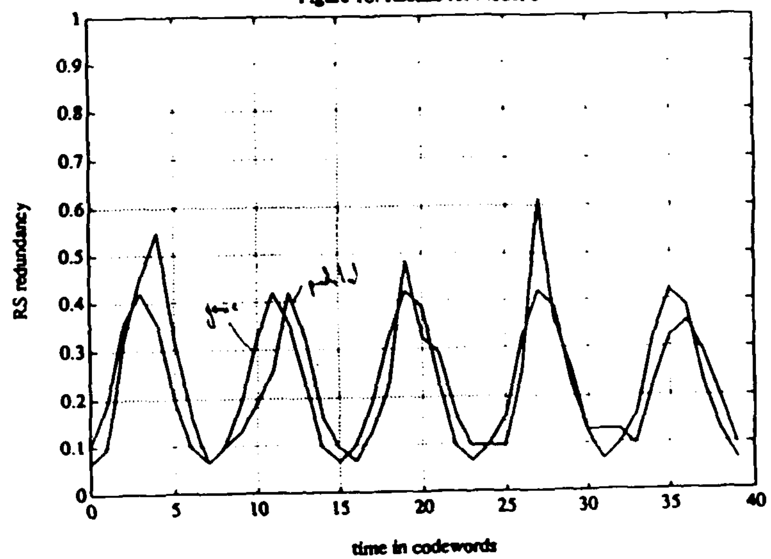


Figure 16: Results for model 6



Channel 7 consists of the same transition matrix as used for Channels 5 and 6, but with the BER parameters varying in an extreme fashion. Figure 17 shows the BER versus time graph. The channel consists of very sharp changes between benevolent and (impossibly) bad channel conditions. Figure 18 shows the effect on the predictions. The genie prediction shows a sharp series of dives from redundancy 1.0 (i.e. the code cannot cope) to about 0.1 and back. The HMM estimation shows virtually the same results except that the best case redundancy tends to be slightly worse than 0.1 and the periods of non-1.0 redundancy tend to be slightly wider; the HMM estimate is based upon observed data and hence periods of improving and declining channel conditions tend to be predicted better than in the single snapshot mechanism of the genie prediction. These results are important for the following reason. At the start of the simulations it was suspected that the HMM estimation process would not be accurate under either of the following conditions:

1. The initial estimate was considerably at variance with the ideal answer;

2. A small number of iterations were used.

Figure 18 shows that the process can adapt and track over large dynamic ranges.

Channel 7b consisted of slowing down the rate of change of conditions in Channel 7. Figure 19 shows the results of the simulation. Over the space of about 11-12 codewords, the genie prediction would increase the redundancy from below 0.1 to the maximum possible. Although necessarily lagging, the HMM estimation process agrees closely with these results. It is interesting that, at the 10-th codeword, the HMM estimate decreases the redundancy.

Figure 20 shows the average BER versus time graph for Channel 8. The added complexity of the four state markov chain clearly does not prevent accurate tracking.

In Channel 8b, the rate of change of noise mechanism was reduced. Figure 22 shows the simulation results. The genie prediction increases the predicted redundancy uniformly. However, although tracking the increase in a broad sense, the HMM estimate shows more divergence from the genie than in the other simulations. The likely explanation is that the channel has a relatively complex noise mechanism (4 states) and that the codeword length of 496 bits is insufficient for the noise samples to present all attributes of the chain to the HMM estimation algorithm. This is discussed in Section 5.

In Channel 9, the transition matrices were varied with time. The time-varying state transition matrix was a mixture of the channel 6 three state matrix and the channel 8 four state matrix. Figures 23 and 24 show that the codes can cope with Channel 9.

5. Discussion of Results and Design Issues

The central question is to the extent that the results indicate than an adaptive coding scheme could be designed around the forward-backward algorithm. The results indicate that the algorithm tracks a channel with a hidden markov chain noise mechanism (of sufficient severity) closely, providing predictions close to the best possible. Moreover, the work of [4-6] indicates that such a noise mechanism is liable to exist on scintillating channels. However there are several points to be made. An adaptive scheme will contain three major components:

1. Codes;

2. Channel Estimation;

Figure 17: BER versus time for model 7

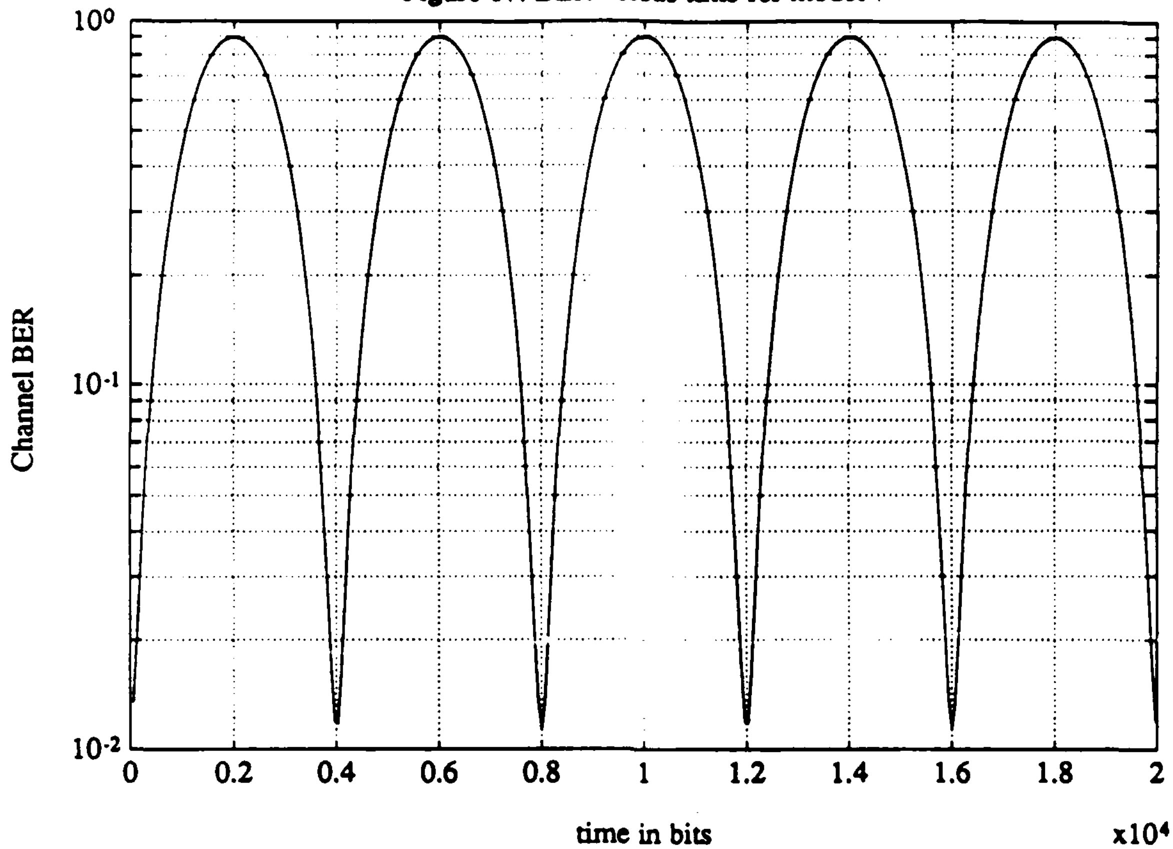


Figure 18: Results for model 7

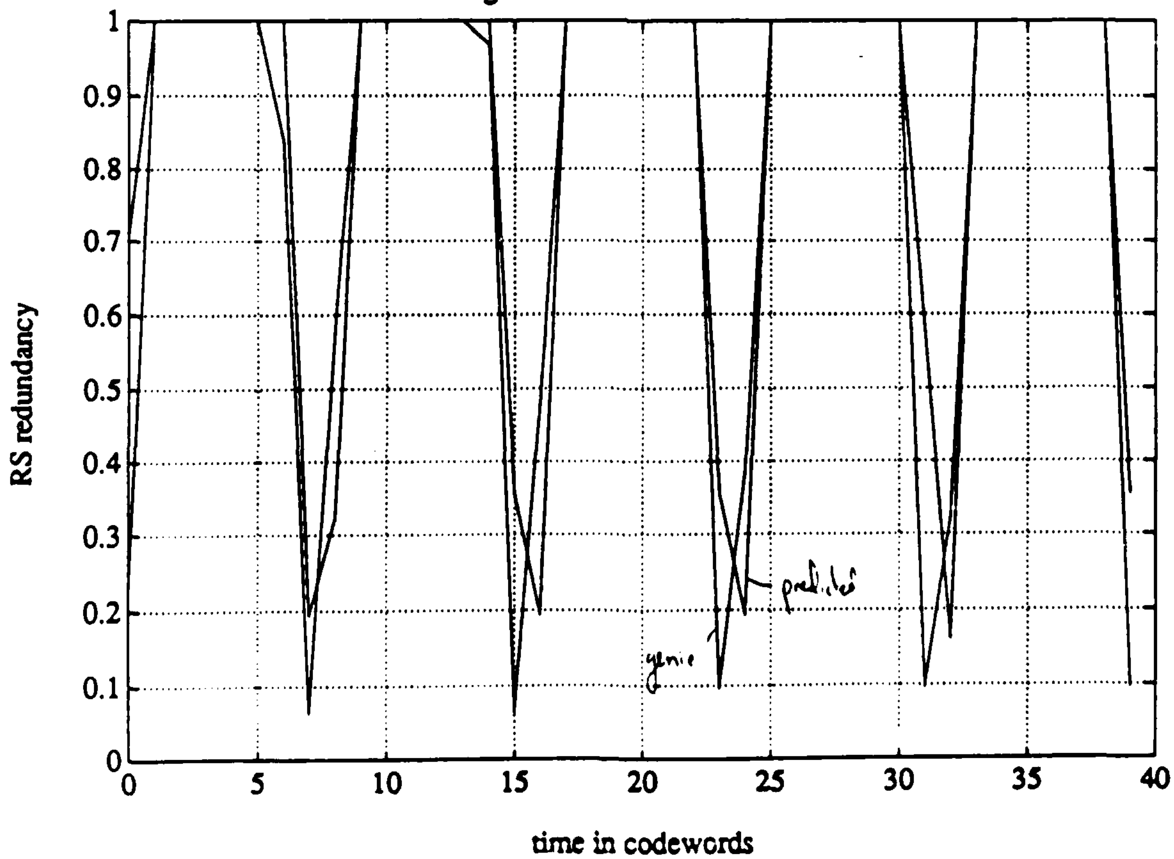


Figure 19: Results for model 7b

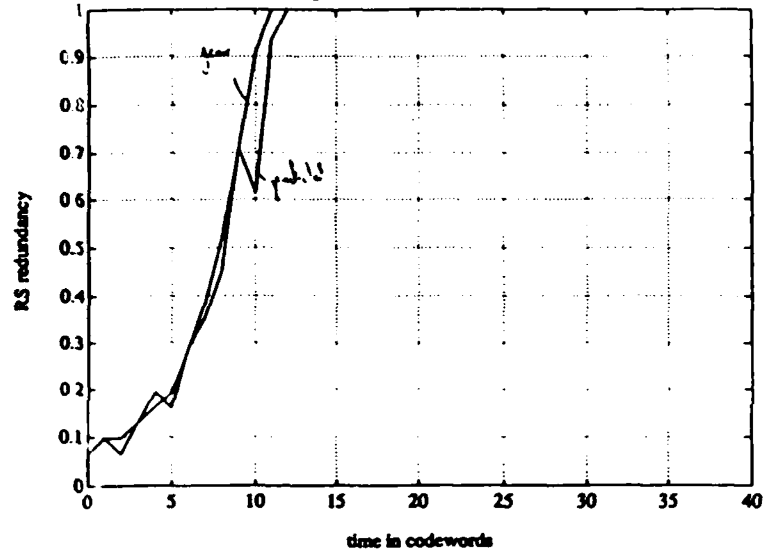


Figure 20: BER versus time for model 8

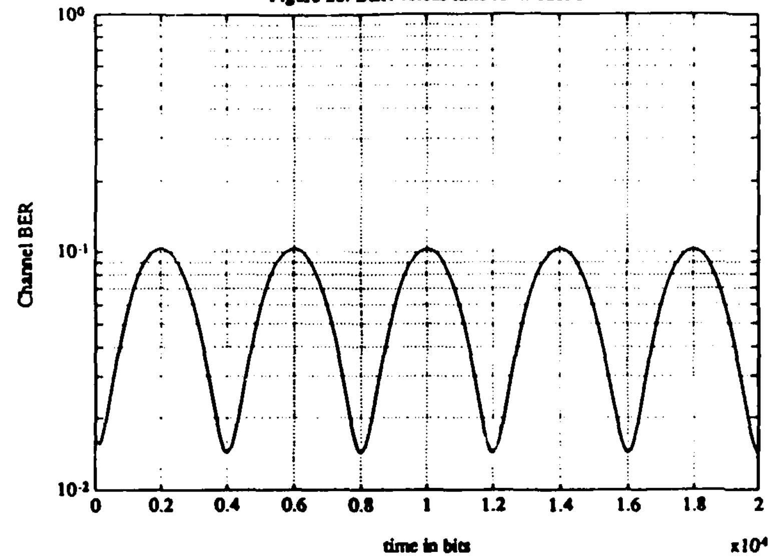


Figure 21: Results for model 8

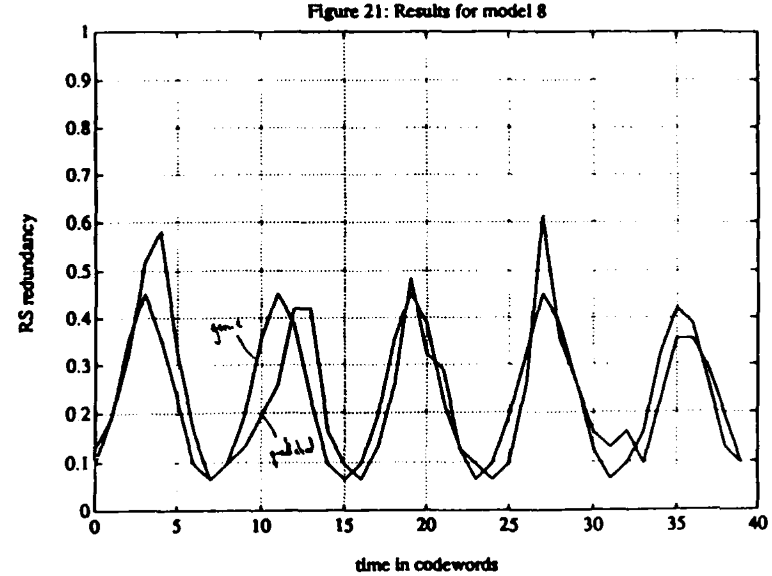


Figure 22: Results for model 8b

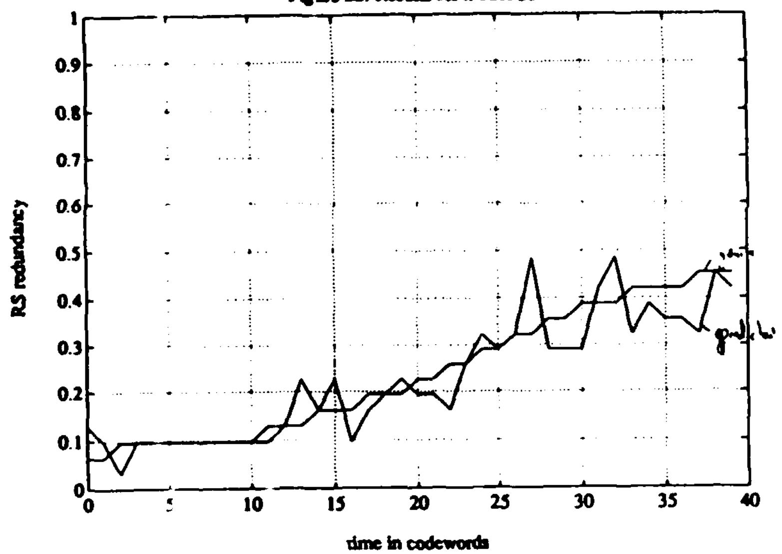


Figure 23: BER versus time for model 9

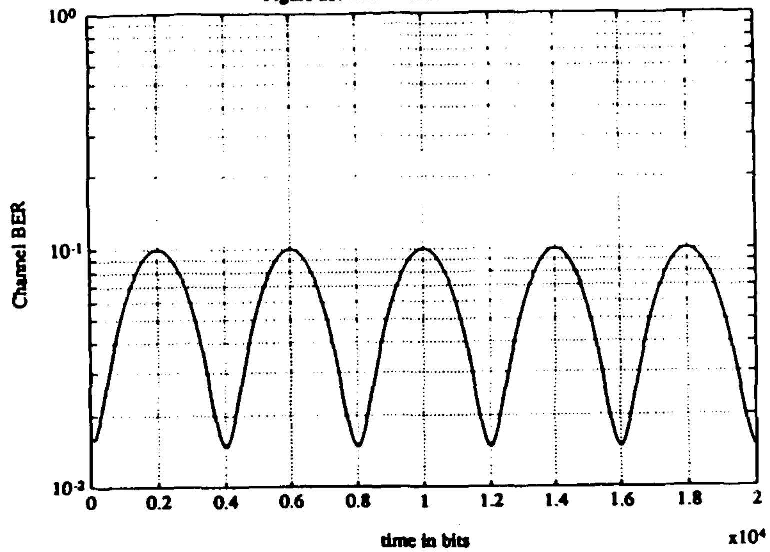
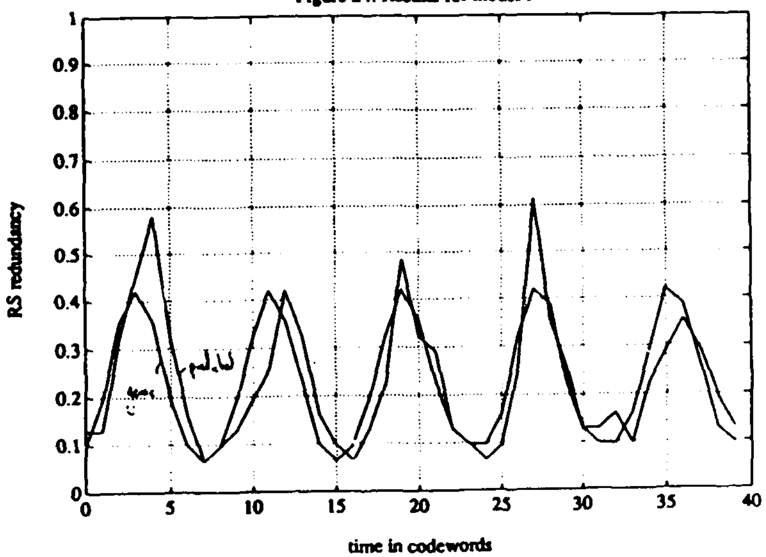


Figure 24: Results for model 9



3. Adaptation Protocols.

The first two areas would appear to be readily designable, codes are readily obtainable and the results of Section 4 indicate an efficient estimation algorithm. However, there are some subtle points. For stationary channels, a worthwhile method to reduce code redundancy and increase performance is to increase the length of the code. For an adaptive scheme this is not necessarily so worthwhile. A longer code length increases the time between code alteration, increases the range of noise conditions that one code must be able to combat, and causes more data to be lost in the event of an inability to decode. Conversely, it is clear that the HMM will more accurately estimate the markov model on longer stretches of data than short. This does not imply that long codes are better; several short codes can be combined to give a longer block. The optimal type of interleaving on an adaptively coded channel is not clear. Note that, if the channel is predominantly non-noisy, long blocks of data are necessary to ensure sufficient amounts of errors.

There are several possible ways of accommodating these points. Perhaps the simplest method is to activate the HMM only when a certain number of errors have been observed (in fact, the simulation software was originally written on this basis). Thus the estimation process is busier during bad error periods than quiet. This could easily be combined with a windowing process whereby each codeword results in noise data that is added to a buffer or shift register. The buffer contains data on the position of the most recent N errors and this data is passed to the estimation process on a once per codeword basis. On a quiet channel, the N errors are obtained from several codewords. On a noisy channel, the data represents less history. This method has the advantage and disadvantage of producing smoothed, averaged models more suited for slowly varying channels than faster varying channels.

For slowly varying channels, a saving in complexity is possible by altering the definition of the markov model. In the simulated model, the HMM produced an output on a once-per-channel-bit basis where the output was a '0' (non-error) or '1' (error) respectively. In the alternative, an HMM produces output on a once-per-error-bit basis where the output is a positive integer equal to the number of bits since the last error. The states of the HMM represent different error rates as before, but the distribution of the output of the state (i.e. an interarrival time) is given by a negative binomial distribution. The effect of this model will be to impose a reduced memory requirement but with a loss in accuracy. (The loss of accuracy is imposed by the estimated state sequence changing only at the error locations. This is not imposed by the simulated once-per-channel-bit model and, hence, the simulated model maximizes the estimate over a larger set and is more accurate.) The once-per-error-bit model will also have the disadvantage of being relatively slow to decrease the redundancy. (This can be ameliorated using a mixture of the two methods.) The once-per-error-bit model seems to be attractive for relatively benign channels, e.g. ISDN links [46].

The protocol design question has not been fully addressed in this study; it would have been presumptive to design a scheme about an unproved estimation process. However, there are some preliminary comments that can be made. In a block code scheme, it is possible to sacrifice one "slot" per block for a control channel. Hence, if we used a

block code of length 32 symbols, we might reserve one symbol for synchronization data and one for code negotiation. (If the noise source is an intelligent jammer, the exact location of these slots within a block should be under cryptographic control.) If these control slots have their data repeated a fixed number of codewords, the resulting slow speed channel can be made more robust (and more redundant!) than the surrounding codewords. Of course, the repetition slows the speed of adaptation commensurately. This had a direct effect on the parameters chosen for the simulation. It is unlikely that a duplex adaptive scheme of the considered type operating on broadly equivalent severe error conditions on both links will be able to adapt the codes within 4-6 codewords of transmission time. (In some scenarios, the assumptions as to turnaround time are pessimistic. Many studies assume a noiseless feedback path and in some "star" networks, this can be approached. The point is that not every duplex link is so fortunate.) If the noise conditions vary slowly in relation to the codewords, it is enough for the protocol to indicate that the redundancy should be kept as presently set, increased by a preset quantity or decreased by a (possibly different) preset quantity. Allowing for acknowledgments, this indicates as little as two bits of data needs to be coded in the control data. For more dramatic changes, the redundancy needs to be changed within increased limits and more control data is necessary. In fact, the coding can be arranged to allow for a limited degree of confusion between the transmitter and receiver! If the codes are correctly designed, a mismatched RS decoder can process a *slightly* different more redundant RS codeword. For example, an RS(31,15) decoder will process an RS(31,14) codeword to produce 15 symbols of data of which 14 are correct and the final symbol is non-data. (In a transform based decoder, this final symbol will be all-zero, allowing for non-activity monitoring to be used as a check for mismatched coding.) Going the other way is more problematic; an RS(31,14) decoder will *not* consider an RS(31,15) codeword to be a valid codeword with one error! (This attribute can be put to use.) Consider a high complexity adaptive scheme in which a receiver has a bank of decoders processing the received channel data in parallel; the most redundant decoder that signals a valid codeword would be trusted as the correct decoder for the transmission. This could be of use when a central receiver with plenty of computing power is listening to many low-complexity receivers transmitting over channels of diverse qualities.) These comments also suggest a practical way of absorbing the variable throughput. It is natural to consider the K data symbols of an RS codeword as K parallel, multiplexed traffic channels. As the codes have their redundancy increased and decreased, the channels are dropped and re-opened accordingly. Given a priority allocation of traffic to the RS data symbols or "channels," the adaptive coding can be easily optimized to get as much of the most important data through a channel as possible.

As a final protocol issue, it should be apparent that the above comments have been written with little concern for the standard OSI model. It is an interesting point as to whether adaptive coding, as defined in this study, is *necessarily* in conflict with the OSI model.

6. Conclusions

On the basis of the simulations, it appears that the forward-backward algorithm can be used for channel modeling. The algorithm appears to track the prevailing con-

ditions closely and quickly. This solves a problem of real-time channel estimation for many classes of problem channels. If this approach is combined with the existing capabilities of forward error correction, adaptive coding schemes become much closer to reality. It must be stressed that existing coding options carry many possible choices of manually selectable codes; the decoding algorithms within a family of codes are very similar. The type of proposed adaptive coding scheme represents the addition of a controlling computer and a protocol scheme to such devices.

Such schemes will adjust the utilized code automatically from a range of possible options in reaction to the channel conditions, maximizing throughput while retaining data integrity. This could enable the technology for a less sophisticated user than at present. This would have important consequences as regards training, efficiency and acceptability of low echelon data communication.

Further work into the definition, implementation, simulation and trial of a full adaptive coding scheme seem justified and desirable.

Acknowledgment

The authors would like to thank Narcisco Tan for his programming of the forwards-backwards algorithm and for his comments as regards the rest of the simulation software. In addition, several conversations with Professor L. R. Welch were extremely useful.

References

1. E.N. Gilbert, "Capacity of a burst-noise channel," *Bell Sys. Tech. J.*, vol. 39, pp. 1253-1260, 1960.

2. E.N. Gilbert, "On the identifiability problem for functions of finite Markov chains," *Ann. Math. Stat.*, vol. 31, pp. 668-697.

3. E.O. Elliot, "A model of the switched telephone network for data communications," *Bell Sys. Tech. J.*, vol. 44, pp. 89-109, January 1965.

4. B.D. Fritchman, "A binary channel characterization using partitioned Markov chains," *IEEE Trans. Inform. Theory*, vol. IT-13, App. 221-236, April 1967.

5. J-P.A. Adoul, B.D. Fritchman, L.M. Kaval, "A critical statistic for channels with memory," *IEEE Trans. Inform. Theory*, vol. IT-18, pp. 133-141, January 1972.

6. P. McManamon, "HF Markov chain models and measured error averages," *IEEE Trans. Commun. Technol.*, vol. COM-18, pp. 201-209, June 1970.

7. L.R. Rabiner and B.H. Juang, "An Introduction to Hidden Markov Models," *IEEE ASSP Magazine*, January 1986.

8. G.D. Forney, Jr., "The Viterbi Algorithm," *Proc. IEEE*, vol. 61, pp. 268-278, 1973.

9. L.R. Bahl, F. Jelinek and R.L. Mercer, "A maximum likelihood approach to continuous speech recognition," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. PAMI-5, no. 2, pp. 179-190, March 1983.

10. D.H. Ballard and J. Sklansky, "A ladder-structured decision tree for recognizing tumors in chest radiographs," *IEEE Trans. Comput.*, vol. C-25, no. 5, pp. 503-513, May 1976.

11. L. Baum and T. Patrie, "Statistical inference for probabilistic functions of finite state markov chains," *Annals of Math. Stat.*, vol. 37, 1966.

12. G. Coleman and H.C. Andrews, "Image segmentation and clustering," *Proc. of IEEE*, vol. 67, no. 5, pp. 773-785, May 1979.

13. D.B. Copper, "Maximum likelihood estimation of markov-process blob boundaries in noisy images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. PAMI-1, no. 4, pp. 372-284, October 1979.

14. D. Garber, "Computational models for texture analysis and texture synthesis," Ph.D. thesis, University of Southern California, May 1981.

15. R.M. Haralick, "Statistical and structural approaches to texture," *IEEE Proceedings*, pp. 786-804, May 1979.

16. B. Julesz, "Visual pattern discrimination," *IRE Trans. on Inform. Theory*, vol. 8, February 1962.

17. R.L. Kashyap and R. Chellappa, "Estimation and choice of neighbors in spatial-interaction models of image," *IEEE Trans. on Inform. Theory*, pp. 60-72, January 1983.

18. K.I. Laws, "Texture image segmentation," Ph.D. thesis, University of Southern California, January 1980.

19. H.Y. Lee and K.E. Price, "Using texture edge information in aerial image segmentation," *USCISG*, vol. 101, pp. 73-79, March 1982.

20. H.Y. Lee, "Extraction of textured regions in aerial imagery," *USCISG*, vol. 102, pp. 53-64, October 1982.

21. L.A. Liporace, "Maximum likelihood estimation for multivariate observations of markov sources," *IEEE Trans. on Inform. Theory*, vol. IT-28, no. 5, pp. 729-734, 1982.

22. B.H. McCormick and S.N. Jayaramamurthy, "A decision theory method for the analysis of texture," *Intl. J. Comput. Inform. Science*, vol. 4, no. 1, pp. 1-37, 1975.

23. U. Montanari, "On the optimal detection of curves in noisy pictures," *Communications of ACM*, vol. 14, pp. 335-345, May 1971.

24. A. Nadas, "Hidden Markov Chains, the Forward-Backward Algorithm, and Initial Statistics," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. ASSP-3, no. 2, pp. 504-505, April 1983.

25. N.E. Nahi and M.H. Jahanshahi, "Image boundary estimation," *IEEE Trans. on Comput.*, vol. C-26, no. 8, pp. 772-781, August 1977.

26. N. Nahi and S. Lopez-Rora, "Estimation detection of object boundaries in noisy images," *IEEE Trans. Automatic Control*, vol. AC-23, no. 5, pp. 834-845, 1978.

27. W.K. Pratt, O.D. Faugeras and A. Gagalowicz, "Applications of stochastic texture field models to image processing," *Proc. of IEEE*, pp. 542-551, 1981.

28. A. Rosenfeld and M. Thurston, "Edge and curve detection for visual scene analysis," *IEEE Trans. Comput.*, vol. C-20, pp. 562-569, May 1971.

29. J.M. Tenenbaum, *et al.*, "Map-guided interpretation of remotely senses imagery," *Proc. Nat. Comput. Conf.*, Anaheim, CA, pp. 391-408, June 1980.

30. C.W. Therrien, "An estimation-theoretic approach to terrain image segmentation," *Computer Vision, Graphics and Image Processing*, vol. 22, pp. 313-326, 1983.

31. J.W. Woods, "Markov image modeling," *IEEE Trans. Automatic Control*, vol. AC-23, pp. 846-850, October 1978.

32. J.K. Baker, "The Dragon System - An Overview," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. ASSP-23, no. 1, pp. 24-29, February 1975.

33. F. Jelinek, "Continuous speech recognition by statistical methods," *Proc. IEEE*, vol. 64, pp. 532-556, April 1976.

34. A.B. Poritz, "Linear predictive hidden markov models and the speech signal," *Proc. ICASSP '82*, Paris, France, pp. 1291-1294, May 1982.

35. H. Bourlard, C.J. Wellekins and H. Ney, "Connected digit recognition using vector quantization," *Proc. ICASSP '84*, San Diego, CA, pp. 26.10.1 - 26.10.4, March 1982.

36. L.R. Rabiner, S.E. Levinson and M.M. Sondhi, "On the application of vector quantization and hidden markov models to speaker-independent, isolated word recognition," *Bell Sys. Tech. J.*, vol. 62, no. 4, pp. 1075-1105, April 1983.

37. B.H. Juang, "On the hidden markov model and dynamic time warping for speech recognition - a unified view," *AT&T Bell Labs Tech. J.*, vol. 63, no. 7, pp. 1213-1243, September 1984.

38. S.E. Levinson, L.R. Rabiner and M.M. Sondhi, "An introduction to the application of the theory of probabilistic functions of a Markov process to automatic speech recognition," *Bell Sys. Tech. J.*, vol. 52, no. 4, part 1, pp. 1035-1074, April 1983.

39. L.E. Baum and J. Eagon, "An inequality with applications to statistical prediction for functions of markov processes and to a model for ecology," *Bull. Amer. Math. Soc.*, vol. 73, pp. 360-363, 1963.

40. L.E. Baum, T. Petrie, G. Soules and N. Weiss, "A maximization technique occurring in the statistical analysis of probabilistic functions of markov chains," *Ann. Math. Statistic.*, vol. 41, pp. 164-171, 1970.

41. L.R. Rabiner, B.H. Juang, S.E. Levinson and M.M. Sondhi, "Recognition of isolated digits using hidden markov models with continuous mixture densities," *AT&T Bell Labs Tech. J.*, vol. 64, no. 3, pp. 1211-1234, July-August 1985.

42. B.H. Juang and L.R. Rabiner, "Mixture autoregressive hidden markov models for speech signals," *IEEE Trans. Acoustics, Speech and Signal Processing*, vol. ASSP-33, no. 6, pp. 1404-1413, December 1985.

43. L.M. Adelman and J. Reeds, "On the cryptanalysis of rotor machines and substitution-permutation networks," *IEEE Trans. Inform. Theory*, vol. IT-28, no. 4, pp. 578-584, July 1982.

44. E.P. Neuberg, "Markov models for phonetic text," *J. Acoust. Soc. Am.*, vol. 50, pp. 116(A), 1971.

45. N-K. Huang, "Markov probabilistic output models for texture-based image segmentation," Ph.D. thesis, University of Southern California, 1983.

46. Y. Yamamoto and T. Wright, "Error performance in evolving digital networks including ISDNs," *IEEE Commun. Mag.*, vol. 27, no. 4, April 1989.

47. E.R. Berlekamp, "The technology of error correcting codes," *Proc. IEEE*, vol. 68, pp. 568-593, May 1980.

48. E.R. Berlekamp, R.E. Peile and S.P. Pope, "The application of error correction to communications," *IEEE Commun. Mag.*, vol. 25, no. 4, April 1987.

49. W.W. Wu, D. Haccoun, R.E. Peile and Y. Hirata, "Coding for satellite communication," *IEEE Selected Areas in Commun.*, vol. SAC-5, no. 4, May 1987.

50. Advanced Train Control Systems Specification 210: Mobile Communications Package, February 1989.

SEYMOUR STEIN: *Systems Perspectives*

I've done a lot of things in my career, but being part of a panel of coding experts is a really new experience. Anyone who knows me probably wondered what I was doing on the program. That was also what I told Rob Peile when he first approached me to serve here.

My view of coding has always been that it's a black art. Antennas are also a black art. As a system designer I wait for the artists to come up and show me what they can do, and I hope I know what to use when I need to do a new system design. I've also learned to be very skeptical when the artists come up with new ways to solve all the world's problems. And Rob said that's exactly the attitude he wanted to close off the panel session, and that's why I'm here. I'm the skeptic.

We also had a deal that I could be a silent skeptic, that I was going to sit up here, not have to give any formal talk until the question period, and then I could join in. Then when I got here we found that Andy Viterbi had cancelled out, and so you get to hear me. The good news about all that (and there's my title chart) is that I have only four charts. [LAUGHTER]

What I'm really going to do is try to talk more generally about adaptation and then make some comments about the possible role of adaptive coding in all this. Both Al Levesque and Rob have anticipated some of my comments, but I'll repeat them anyway.

There are two primary reasons for adaptation in the radio channel. [CHART #1] One reason for adaptations are the kinds of problems we find in fading multipath channels or when we have moving terminals, more general than only HF channels. There are obviously changes in channel gain. Typically there are short-term changes that occur in very short time scales like seconds, along with longer term changes. There are also long-term changes in the nature of the channel transfer function as the character of the multipath somehow changes. I am a very proper skeptic, and you'll see why, as to whether adaptation to the short-term changes is really practical. For a longer term change, adaptation may be practical but needs to be put into perspective.

The other kind of change that certainly the Army customer here or any military spon-

REASONS FOR ADAPTATION

- SIGNAL CHANGES:
  - CHANGES IN CHANNEL GAIN
    - SHORT TERM (E.G. RAYLEIGH FADING)
    - LONG TERM (E.G. DIURNAL)
  - CHANGES IN CHANNEL DISTORTION
    - MULTIPATH CHARACTER
- NOISE CHANGES:
  - BENIGN
    - INTERFERERS
    - IMPULSES
  - HOSTILE
    - JAMMING

CHART #1

sor is interested in is the kind of change that you might think of in response to jamming. These are applications where interferers that weren't there before suddenly show up and you have to try to cope with them – that is another kind of change. These are, in general, the reasons people adapt.

Another issue on adaptation that you have to recognize – and again, especially in the military environment – is that we live in a world where some messages are very short. Adaptation in reality may be what you do in order to try to find a way to get the message through in the first place. You don't have very much time for fancy adaptation like changing codes because if you've gotten through at all, that's it. So that in the sense it's been talked about here, adaptation really refers to more continuous kind of links.

Now that word "adaptation" or "adaptive operation" has been applied in all kinds of places, and I'm sure this is not a complete list [CHART #2] but let me just list them. One of the most important, most significant kinds of adaptation that has nothing to do with coding is that of using relay networks with alternate routing. It's the kind of thing that Al Levesque referred to as the probable future of HF, but it applies to much more than HF, and I'd like to make an aside here that has little to do with the topic.

While coming up here I happened to read a trade journal in which some Army general proudly announced that the Army is now within five years of completing its procurement of a system called TRITAC which has been in development and procurement for twenty years. In 1995 the Army, as he said, is finally going to have a battlefield telephony system which is fully digital, fully encrypted; it's going to do all kinds of great things. The more important comment was: now that they can see an end to the billions of dollars they're spending on that system, it's time to start thinking about the next generation. If anybody here is looking for new problems to conquer, I would very much recommend thinking about how to build relay networks that operate on the battlefield. I really think that down to even the lowest echelon, the days of omnidirectional communication in an electronic warfare environment are numbered. We're going to have to go up in frequency, finding ways to use directional transmission. That's an aside, and as I said it has nothing to do with HF or coding.

Another kind of adaptation is frequency change. You find interference or you find you're not getting through at one frequency, you try another frequency. Some of the newer HF systems in fact do exactly that and they have been delighting the users because they work. The traditional HF operation was relatively fixed frequency. The so-called "selective calling" systems try multiple frequencies until they find one that gets through.

Another kind of adaptation, transmit power level control, is one that is of great interest to a lot of people, especially people trying to design covert communication systems who want to operate at the lowest power level. There is a slight connection with coding but probably not with adaptive coding.

Receiver processing in many forms has also been called adaptation. One example is the dynamic operation of an equalizer – we call them adaptive equalizers. They don't require two-way communications, however. I call this a kind of implicit adaptation. That is, you build something like a Rake receiver, a diversity combiner or an equalizer that simply operates within a limited set of variations and attempts to give you the best throughput in the light of those variations.

TYPES OF ADAPTATION

- ALTERNATE ROUTING THROUGH RELAY NETWORK
  - INCLUDING ADAPTIVE RADIO PATH SELECTION
- FREQUENCY CHANGE
- TRANSMIT POWER LEVEL CONTROL
- RECEIVER PROCESSING
  - FILTERING TO MINIMIZE NOISE
  - MATCHING TO SIGNAL VARIATIONS (DIVERSITY, INCL. RAKE)
- VARY BANDWIDTH (FIXED THRUPUT RATE)
  - SPECTRUM SPREADING
  - CHANGE CODE REDUNDANCY (CODE RATE)
- CHANGE THRUPUT RATE (FIXED BANDWIDTH)
  - CHANGE MODULATION

    SPREADING

    SIGNAL CONSTELLATION
  - CHANGE CODE REDUNDANCY
  - IMPLICIT VARIATION

    ERROR DETECTION (POS. ACK) + REPEATS (INCLUDES CODE COMBINING)

CHART #2

The two areas that I can see that relate to coding are: 1) the possibility of varying the bandwidth, which again is not the subject of this panel; the other is 2) changing the throughput rate, which is what I think people mean by adaptive coding. Varying the bandwidth, while not the subject of this panel, is a very respectable military notion because that's what spectrum spreading is all about. By the way, for many reasons, I like to regard spectrum spreading as simply another kind of code. It is in effect a low rate code if you think about the bandwidth that's being used. In principle one could alter the code redundancy so as to use no more frequency than you really need at any time. I don't know that anyone has proposed that for spreading and, again, I'm not going to say much more about it.

Finally, changing the throughput rate is what I believe we are talking about as adaptive coding. There are three ways, all of which have been mentioned, I believe, for changing the throughput rate. One is, obviously, to change the modulation that you're using. Again within a fixed bandwidth you can vary the amount of spectrum spreading you're using. A classic statement about how a military equipment might be designed to go from a benign to a jammed environment is, "Gee, we'll lower the data rate and just use the bandwidth that we've made available to get processing gain against the jamming." A less sophisticated thing would be to just to think of altering signal constellations to change the data throughput rate. If your object is to work at lower or higher E_b/N_0, a change in the constellation will do that. Changing code redundancy, again, also does exactly that. It changes in effect the E_b/N_0.

And finally there is what I called here an "implicit variation." You can use error detection which in effect uses up part of your transmission time when you're not able to get through. It effectively lowers the average data rate because you're doing repeats instead of sending new data.

Now let's focus on the non-implicit forms of adaptation, the forms of adaptation that require two-way coordination, where we require the transmitter to do something in response to information sent from the receiver and also require the receiver to know what the transmitter is doing. I've listed here [CHART #3] some of the classic problems that are encountered any time anybody has tried to investigate such a system. One is the very problem of trying to estimate exactly what the problem is that you're trying to cure because, obviously, the effectiveness with which you cure it depends on how good you are in estimating what the problem is. The other issue that arises, and perhaps the strongest reason for being skeptical about using any kind of adaptation to cope with short-term fading, is that in general the estimation process requires time, smoothing time. Very often that smoothing time turns out to be longer than the time within which the channel stays stationary. The coordination you require between terminals requires a two-way link, and there are various ways to perhaps make that return service link more robust than the comm link. But in doing that, again, you give up channel capacity in that reverse direction in some way. A bigger problem when people have looked at feedback of that type is that the delays and the errors do lead to potential system level instabilities.

Another aspect of adaptive coding that would begin to trouble me, especially after some of the numbers I've heard, is the very fact that in thinking of trying to build very robust codes you typically end up talking

SYSTEM ISSUES IN ADAPTATION
(NON-IMPLICIT)

- MEASUREMENT/ESTIMATION PROCESS

  ERRORS

  SMOOTHING TIME VS. DYNAMICS OF PROBLEM

- NEED FOR COORDINATION BETWEEN TERMINALS

  REQUIRES OPERABLE TWO-WAY LINK

  POTENTIAL INSTABILITIES DUE TO DELAYS

- CODING AND DECODING DELAYS (INCLUDING INTERLEAVING)

- COMPLEXITY LIMITS IN IMPLEMENTING ADAPTATION

CHART #3

ADAPTIVE CODING ISSUES

1. DS SPREADING, WITH CHIP RATE VARIABLE (ANTI-JAM)
   SYMBOL MATCHED FILTERS FAVORS USE OF "RANDOM" SEQUENCES

IS THERE A BETTER WAY TO SELECT SPREADING CODES?

2. VARYING CODE RATE WITHIN FIXED BANDWIDTH

$$SNR_{REQ} = (E_b/N_0)R$$

- CODE SELECTION CAN CONTROL (E_b/N_0) FROM 12-15 dB (UNCODED) DOWN TO 1.5 dB (3 dB PRACTICAL LIMIT?)

- CAN VARY R DIRECTLY VS. AVAILABLE SNR WITHOUT ADAPTIVE CODING

CAN BENEFITS OF ADAPTIVE CODE SELECTION JUSTIFY THE COMPLEXITY/COST?

CHART #4

about very long codes, and long codes and interleaving introduce delay. Again, there are many applications in which delay is almost a no-no as far as the user is concerned. Somebody being shot at doesn't want to be told that he's getting better communication if he can only wait. The other problem with those delays is they foreclose on very rapid or real time adaptation to very short-term changes.

Finally the issue always in any kind of adaptation will be the cost in trying to estimate and understand the complexity limits that are tolerable. That cost changes with time but it's always there.

Finally I'll get to perhaps two questions that I might challenge the adaptive coding people with [CHART #4]. One is with respect to viewing spread spectrum as a kind of low rate coding. We always think of selecting spreading codes for randomness and we tend to think of the codes in some sense as being very random selections in a very large space. It occurs to me that I had never thought before to even ask the question: We have such a large space available. If we know something about the problem we're coping with, for example the nature of the jamming, might there be subspaces that might still be very large and give us a lot of the advantages we're looking for but that might operate very effectively against some jammer *that itself was not* blanketing the entire signal space? The challenge would be to figure out a way to convert information like "this is a tone jammer that's in this part of my spectrum" into information on the design of a code.

With respect to the kind of adaptive coding that this panel has been looking at, I have a very naive view of that, probably overly naive. Whatever the code, we describe its operation ultimately in terms of the E_b/N_0 required per data bit, and therefore always the signal-to-noise ratio that's required to get the performance we want is simply a product of that E_b/N_0 and the throughput rate that we're getting. One of the practical problems is that the range of E_b/N_0 that we can do something about by coding is not very great. If we go from a totally uncoded situation to perhaps the computation limit, we're talking about a range in SNR from about 12-15 dB *down to perhaps* 1.5 dB. *This would be* for an unfaded channel. For a faded channel those numbers might be from 20 dB down to perhaps 7 or 8 dB. When you think about problems like jamming, or on HF propagation paths dropping out on you, which are the kinds of problems that you might think you want to adapt to, it's not very clear that that total range of 10 dB is really going to buy you very much. Particularly we never operate uncoded these days, most designs include some simple codes to begin with, so that coding adaptation, it seems to me, is really unlikely to offer much in the way of improvement in the ability to deal with significant changes in SNR.

The final comment here, very simply, is a complexity issue. If I can find a way to coordinate, obviously I can vary the code rate just by a simple modulation change, or by repeating bits in a very simple and uncomplex way without any need for adaptive coding. So my question obviously comes back again: Can the benefits be in some way shown to justify what I take to be the probable complexity and cost of more sophisticated adaptive coding? Of course that's going to give Rob and a few others several years of further work in order to try to justify all that. Thank you.

PEILE: OK, I think we'll take a break for 10 or 15 minutes and then we'll have a discussion session. Ten minutes, OK! [PAUSE]

Note that we have not talked about adap-

tive FEC on a network level, where FEC gets involved in adaptive routing strategies. Work does exist in that area. Missing packet problems and whatever hasn't been covered; we haven't covered everything.

LLOYD WELCH: In your cyclostationary simulations, where the code rate went up and down, a thought that occurred to me: Does it go down fast enough to prevent any failure to decode indicators?

PEILE: Well, that's a good question. My simulations produced three columns of data. The column I did not show gave the redundancy you should have used if you had known what the noise was going to be. The actual noise samples differ, of course, from the noise model. Generally speaking, if you could decode at all from the observed noise, you could decode with the predicted code. In other words, the predicted code was able to cope when things were capable of being coped with.

I have a few examples where the noise samples scared the hidden Markov model into stating that it could not cope, even though the next noise samples were not too bad. When you looked at what was going on, in fact you could cope even on these examples. That's because my predictions tried to give a 10^{-4} error rate. In practice you would simply pin the redundancy to a pre-set maximum. The predictions were pessimistic. It tended to overshoot the redundancy rather than undershoot. The predicted redundancy tended to come down and get it right. The valleys were OK; the peaks were what shot up.

PAUL SASS: I'd like to just amplify on a comment that Seymour Stein made regarding this 1995 date. In fact it's accurate. In 1995 the Army will not only have finished purchasing its TRITAC system, which is the echelon above Core system, but will also have finished buying the tactical communication system below Core – the mobile subscriber equipment, the combat net radio and our data distribution system. I'll be amplifying on this a little more tomorrow.

I think our challenge in this community is to target our research products now so that come 1995 we can kind of stage an attack on improving these systems. We're not going to come in with a new system in 1995 and replace this investment that the Army has just spent five or ten years on. We're going to have to improve the coexistence of these systems, we're going to have to improve the survivability of these systems, we're going to have to improve the connectivity of these systems. That's what I'd like to see the research in this community start targeting, so come 1995 we can really apply these new concepts now.

MIKE PURSLEY: *I'd just like to mention* two areas of application of adaptive coding that haven't come up here, although Rob briefly alluded to the mobile tactical packet radio network. There certainly we have a very important application of adaptive coding. If you envision mobile terminals and possibly mobile jammers and other sources of interference, then what you find is the link condition is varied tremendously throughout the network. The forward error control coding on those links ought to account for what the conditions are to the best that it can be informed. The information there generally is passed around very much in the same way and may be part of the information that must be passed around for adaptive routing and other adaptive protocols as well. Adaptive forward error correction certainly fits in very well in the overall scheme of the radio network and is one of the elements of an adaptive protocol sweep that needs to be developed and understood a great deal more.

The other area that I'm familiar with that hasn't been touched on is in the meteor burst communication area. There are two uses for adaptive coding. Most of the trails that appear are these so-called underdense trails where you have an exponential decay in the received signal power level as a function of time. You'd like to do some kind of adaptive coding there to maximize the amount of information you can get over the channel during the lifetime of the trail. It turns out that you have to estimate not only the receive power levels, but if you believe the exponential model you also have to estimate the decay rate and feed that into your adaptive coding. There's also the intertrail adaptivity that's important. Some of the trails are actually so-called overdense where you have stronger signal power levels for a longer period of time. Of course there you'd like to use a very low rate code because you have a very clean signal, in that situation; you'd like to get high throughput – high rate code, I'm sorry, I said that wrongly.

STEIN: On those channels, the mobile channels and the meteor ones, do you know enough about the variations you're trying to contend with to believe that changes in code rate will really do the job? That's the question I was trying to ask.

PURSLEY: OK, well let me talk first about the mobile radio network. Again the information comes from the interference conditions that is observed by your neighbors, and it gets passed around in the control and also in the data packets. Just to the same extent that you can do adaptivity with any of the protocols, you can do adaptivity with the rate of the code. And yes, in that case, I think you do know quite bit. We've done some simulations on mobile partial band jammers and frequency hop radio network, for example, that show quite a bit of improvement that can be obtained through the use of adaptive coding and adaptive protocols, and more generally a wide range of adaptive protocols – adaptive transmission, adaptive forwarding, adaptive routing, adaptive error control coding.

In the meteor burst channel the problem, as I mentioned, is one of estimating the decay rate on the channel. And that's not an easy problem; I don't want to minimize that. It looks like what you probably want to do is to start out with a pre-selected fixed rate code until your estimates are reliable enough to tell you enough about the channel which we begin to adapt intelligently. I think it remains to be seen how much improvement can be obtained there.

PEILE: Thank you, Mike. I agree we haven't looked at that.

BILL LINDSEY: I'd like to make some comments. In order to address this community I have to adapt myself – sometimes I'm confused as to what I really am: whether I'm an analyst, a systems engineer or a professor.

What I want to amplify exactly is in some respect on what Seymour was saying and also what Paul has to say. First of all, it is a very important issue, how do we transition existing complex high-cost systems into networks that are currently planned and currently operating? It is not a simple matter to exchange codex and modems and operating procedures, just because we as researchers have discovered ways to adapt algorithms and modify them in real times. Certain environments such as the military channel – for instance, once you start putting closed loop type algorithms in a system that can be jammed, especially from one end or the other, you have problems with complex issues regarding how to mitigate against this kind of thing. It is a

tactic that we use in AJ systems to attempt to design open loop algorithms, as opposed to closed loop techniques such as phase locked devices and frequently there's a reason .... One of the drivers we do not say go to coherent systems aside from the issue of hopping bandwidth and what not.

The notion of interoperability amongst satellites and users has been around for a long time. It's sort of a requirement that the military has laid upon the community for a period of time and people have talked and were fighting about waveform standards, waveform designs, and they have not reached a solution with regard to that. The problem is that there are different environments, there are different frequency bands, the data types are different, voice in some case, 75 bits in some cases, and he'd like video data in some cases. So throttling algorithms and data rates and codex and modems is a very risky, risky situation as far as I'm concerned.

Now that discussion that I'm making is probably addressed from the satellite communication point of view. I suspect strongly that it doesn't hold when you talk about telephone channel where I think bandwidth is fixed. Bandwidth is fixed and hence in an AJ channel, for instance, you have bandwidth to play with although there becomes, at this point in time, reasons to worry about adjacent channel interference even though you have quite a large bandwidth to use.

It is that the requirements of ... in an AJ system it is my belief that the system design as far as protection in the AJ environment is that low E_b/N_0 is the requirement. Some of these coding techniques that we're talking about will force you to work with higher E_b/N_0's, hence the jamming immunity will be suffered in some sense.

Finally I think Seymour hit upon it, although he didn't mention this specifically: In channels which scintillate, it turns out even at the present time, the memory that is required in an interleaver at the data rates we'd like to push through channels is not sufficient from a technology point of view. Hence channel memory is an issue. That is to say we as coders, or you as coders and me as critique at the moment, tend to dismiss the notion of channel with memory. You always start out by assuming a memoryless channel and hence you go from there by sticking in an interleaver. That's valid insofar as it goes. I think Bob Peile's work that he's doing is a correct direction with respect to adaptive coding in the sense that he is attempting to accommodate some residual memory that one may not be able to overcome by using long interleaver at the transmit side. The issues of delay I think Seymour brought out very well. I brought up jammer issues and closed loops, low E_b/N_0 designs.

In the HF channel itself, although I'm not an expert, I tend to think when you start networking, that one has to be careful and not have to sit back and say, "In order to control this network you have to postulate a separate control channel and control from a separate site," which would be one possible way of doing this. There are a lot of system issues here that should be taken into consideration when one attempts to direct his research toward new concepts – what should I say – in coding.

I think that those are my comments. I believe that Bob Peile's work is very interesting – not to say that the other isn't interesting – with respect to the telephone channel which is a channel I've not really worked, I've just read about it a little bit. The meteor channel is a channel that's very interesting but I think he hit upon what the issues are there.

PEILE: I'd just like to say a few comments both to Seymour Stein, who I deliberately put on the panel to raise these questions, and Bill's critique. Firstly, I think we ought to make clear what we don't understand, or at least, I don't understand. Let me quote Andreas if he's still here. The other day I thought he hit the nail on the head when he said that we might partially understand some techniques like code combining, but there are still some very tricky questions, and we do not understand the full implication of Type 1 and Type 2 hybrid ARQs; we're still researching. We might know how to implement the codes these days, but we don't quite know what to do with them in the fields of ARQ. Again adaptive FEC is new, it's easy to criticize it.

The point about any error correction decoder is that, to do its job, it has to locate the error information. You either throw that information away or you can exploit it in some other part of the system. Most adaptive schemes involving forward error correction say, "I'm not going to throw away the error location information; I can use it." Once you start to relate that information to the task confronting other parts of adaptive systems, then you should see coding as a component for a full adaptive system.

I'd like to mention two things, in particular, to open up a few ideas. One is, whenever I've talked to an HF engineer and talked about coding, he said, "Nah, all you do is change the frequency to a better one." And it's true, I mean, I've played with HF radios and you do that. On the other hand you can turn that answer on its head and reply, "OK, if real time frequency management really is the key to HF or one of the keys to HF, why cannot forward error correction be part of that?" Suppose you have a set of possible frequencies that you have classified as "good" or "bad". You select several "good" and "bad" channels, code over all of them and transmit. At the decoder, if the error correction works, you will learn the accuracy of your "good" and "bad" assessments, without necessarily losing data quality. This can become part of an adaptive hop set scheme.

I'm not saying adaptive coding is the cure-all. All I'm saying is that once you start to think of a decoder as the source of information on the system you're operating, then the critique is not whether or not coding is worthwhile but whether or not you can use the information you've got anyway. That's the slightly different perspective, I've got it.

PURSLEY: Mike Pursley, again. Paul mentioned the equipment that we're going to have to work with and certainly one of the things that we're going to have available, I believe, is the SINCGARS radio. The nice thing about that application, at least as far as I can follow Bill's list of concerns, I'm not sure that any of those things causes a problem in the SINCGARS application. The SINCGARS packet overlay ... therefore to make it into a packet radio network, we'll incorporate Reed-Solomon coding, and we have looked at adaptivity and we've looked at the kinds of problems that Bill is concerned about. Now I won't say that we don't have those concerns for other radios – direct sequence radios or non-spread radios – I certainly think we do.

PEILE: I would say that the frequency hopping radio with a constant band jammer is an application where coding gains can be very large. This is a classical case where coding can give you gains out of all proportion to an E_b/N_0 measurement, simply because how jammed a jammed channel is might not matter. There are limits to E_b/N_0.

STEIN: Can I get one in quick or ....

LINDSEY: Oh, excuse me, go ahead ....

STEIN: Is it live here? This one is live ....

PEILE: What if I say ....

STEIN: No one is doubting the importance of coding in radios. Coding has made a tremendous difference right now in any kind of HF modem design because coding has finally allowed HF communicators to get reasonable diversity at reasonable cost. The only questions that I was asking was about whether you want to start playing games by changing the codes in midstream. It's the adaption part that I'm being skeptical about. As far as the effectiveness of codes in any of these channels, there should be no argument whatsoever. Tremendous!

PEILE: Can you have cross-channel coding?

STEIN: Sure.

LINDSEY: Let me make another comment with respect to Bob Peile's partial band argument. If it is that you're designing an AJ modem, I'd like to make the comment that one of the things you try to do is, of course, use matched filter theory. And in using matched filter theory you usually say that the noise is white, you've got the best thing that you can do and we understand what detection says in terms of optimum reception there. In designing an AJ modem against various jammers – and let's just take the partial band jammer – one of the things that you do in the processing in the modem if you're smart, you want to win a contract, you figure out a way to mitigate that jammer and turn him into a white noise jammer. Hence the algorithm is optimum in the face of white noise, so you should go back to the code that worked in that channel.

PEILE: OK, I'd like to give a reference. I remember that Milstein gave a good paper on coding for hoppers at Brighton which I remember very well. I'm used to the problem of coding for hoppers and not necessarily from jammers, but just from interferers. I remember seeing some statistics of channel occupancy on VHF radio in Germany. Every Army in the vicinity was using the same band and each frequency was used, considerably overused. It's not necessarily a deliberate partial band jammer. In general, if the noise is not white, why not exploit the phenomena?

STEIN: Well, again, let me ... for some reason people seem to be astounded at the notion that coding is going to make frequency hopping work. The correct statement is that nobody has every designed a frequency hopping system without starting out with the idea of having to introduce coding, because you know you're going to take hits. So coding has been an integral part of frequency hopping for thirty years, for as long as hoppers have been designed. The only thing that's happened recently is that codes have gotten more sophisticated.

LINDSEY: My comment was similar to what Seymour has to say. I think that I should keep my mouth shut now, I'm among friends, and talk with Bob off-line. I think it's very interesting and it's not that I'm against adaptive coding at all, and I don't want that to leak out with respect to these minutes. It is the notion of [LAUGHTER] ... it does have to do with existing codes and algorithms. It's very complex. You have several terminals fielded and you attempted to coordinate them to serve all their algorithms, and what one guy needs is most likely what the other guy doesn't need. From a point-to-point point of view, from say Point A to Point B, the notion of adaptive coding as I've heard here today has certainly been presented from a proper perspective. But looking at it

from systems issues, networking issues, I'm very skeptical about adaptive coding. I'm not skeptical about adaptive data rates, lowering the throughput; that can be done in order to achieve some type of diversity, to mitigate channel interference. So, Bob, you and I can have some fun maybe next week continuing this discussion.

WELCH: OK, I'd like to ask Mike Pursley a question. Is the time duration of these meteor bursts such that you can do this adaptive coding? I mean, after all, the transmitter has to decode some codewords, estimate what the channel is doing, feed that information back to the receiver. Do you have enough time to do all that in a fair meteor burst?

PURSLEY: Well, what happens is you don't know when you start how long this trail is going to last. If it's a very short trail, probably you're not going to adapt anything; you'll stay with your fixed rate code. On the other hand if you happen to get a long trail, you can milk it for all it's worth, especially if you have duplex operation which is what we're primarily envisioning here. You can milk more out of the tails when the power level is beginning to drop off – you know, you can drop the rate of your code and get some additional throughput on that trail. When you begin you don't know what you're going to get out of it. Maybe you won't have time to adapt at all and you'll stay with the fixed rate code. If the trail lasts long enough then you can begin to adapt and take advantage of the longer lasting trail. But I think the key question still is: Can you estimate the decay rate and how accurately do you have to estimate the decay rate to get reasonable adaptation? We don't know that yet.

LEVESQUE: Al Levesque. I'd like to add a comment to what Mike has just said. As I recall, I think it was the last MILCOM session in San Diego, Ken Brayer at MITRE presented some interesting results that were based on, as I recall, error data collected on meteor burst channels. He evaluated the use of rather simple codes. I think he looked at simple error detection on short blocks, i.e. blocks that were typically shorter than the meteor trails. In addition, he looked at some code selections that did a small amount of error correction in addition to residual error detection.

Those results, while I don't remember all the details, seem to make a good case for ARQ with block lengths that were expected to be relatively short in contrast with the length of the meteor trails. Then you get to the end of the trail. You simply can't get the blocks through any more and, at some point, even repetition doesn't do you any good because the channel disappears. My question is: Why wouldn't you view that as a good approach to dealing with this unknown exponential decay? I'm not sure that I understand the need for knowing in detail what the exponential decay of the trail is going to be.

PURSLEY: I'm not sure I really understand your question. Certainly if we know that the trail is decaying, which for underdense trails they will, we clearly should have more redundancy in the code to correct more errors. I'm not quite sure what you're asking me in addition to that.

LEVESQUE: Well, as I recall, those results showed that when the trail was strong there were no errors to correct. There was a transition period as the trail began to die where a small amount of error correction seemed to do some good. But if you believe Brayer's results, there was very little value in adding much error correction beyond just a few errors per block, because beyond that point the channel disappeared anyway. In

other words, a very simple scheme in which you used ARQ with blocks that were significantly shorter than the expected length of the trail seemed to do the job very well. In the early part of it when the trail was well established you were getting the blocks through with no detected errors. In that transition region there was some value to doing a small amount of error correction, but then after that the trail was disappearing anyway.

PURSLEY: I'm not quite sure how we're differing here. We certainly start out at the very high rate code and then we decrease the rate as we get into this transition region that you're talking about. We can operate – what I should say is we can get good performance until we lose lock on the signal or something like that. But we can continue to drop the rate and get better and better performance out of it. We presented a paper in the same session, by the way, that showed some of these optimal code configurations for packets as a function of the decay rate on the trail.

WELCH: Lloyd Welch, again. I think I might be leaning towards Seymour's view with respect to the meteor trails. If there really is a decay, then a simple approach is, once you've got a code that works, when the signal drops by 3 dB, I will decrease my baud rate by 3 dB. I maintain the same E_b/N_0 off of the reflection.

PURSLEY: Actually we've looked at a combination of adjusting the symbol duration and the rate of the code. And yes, depending on the type of radio you're dealing with, you may be able to accomplish everything you want simply by changing the symbol duration. If it's a fixed rate radio you may not be able to do that.

LINDSEY: This is not technical, it's not picking on anybody. [LAUGHTER]

Paul Sass made an interesting comment. His comment was that what we should be doing as researchers is to provide him or the Army or whoever this customer may be, "with techniques which can be incorporated systematically post-1995 or whatever year that is." Well, in order for this community – I'm talking about the research community – to help meet that requirement and help give you and give the country the greatest help in this direction, there should be a transfer of information back the other way to this community, which we're lacking. This session, in my opinion, manifests that clearly. We, as a community of researchers at the universities, specifically do not know in general what the requirements are relative to NASA communications, military communications – those things that are evolving with time, forcing requirements on system design engineers. So in order for researchers to guide themselves and try to target things that will solve problems, say post-1995, we need to know what the problem is and how to transition to it, so that we may transition there. And so it is that the problems have to be defined by someone other than the guy who is doing the research, it sounds to me like. That's not the way it's been done. Usually the innovative and creative knowledge bases are developed at the university level and, of course, in industry too. So that was a comment that I thought might be interesting.

WELCH: Bill, you're an optimist, if you think the customer can tell you what he needs ... [LAUGHTER] As a designer of error correcting codes, the most we could hope for out of a customer is a bit error rate. I mean, nothing about the structure of the error symbols or anything else.

PEILE: Absolutely nothing.

WELCH: Absolutely nothing. They do not know anything about how to describe the

error phenomena on a channel except maybe a bit error rate.

PEILE: This is an important point. Lloyd and I have worked together and we do get customers who'll just say it's got errors and ....

WELCH: It's bad.

PEILE: ... it's very bad ....

WELCH: They want to fix it!

PEILE: ... the details are your problem.

LINDSEY: No, I believe you and I understand it. Frequently one of your major jobs in Phase 1 is to define the customer's problem and what his requirements are. I understand that. But the point I'm making is it still is to motivate, if I could or if I can or if we can, a flow of information back from the community who procures and makes acquisitions of such complex systems. And by systems I mean systems composed of subsystems which involve codes, modems, upconverters, antennas, etc. that go in to make up a communication system.

UNINTERPRETABLE DISCUSSION ONGOING IN BACKGROUND

WEBER: Gaylord has been pretty quiet through this whole thing ....

GAYLORD HUTH: Yeah, I've been really quiet the whole time. Let's see, one of the things that I wanted to talk about in terms of adaptive coding is different than what we've talked about so far. It was mentioned here a couple times that as we go in the system with the Army we're going to be talking about maybe higher frequencies, millimeter waves and so on. And maybe you're not jammed any more. I think we've played the game with jamming pretty heavily. I think we understand a lot on the link level – you know, how to protect in terms of coding and spread spectrum and so on. I don't know whether we know that at the network level, and I think there's a lot still to be done there.

The other thing that's interesting is you go to millimeter waves, and especially if you get really high frequencies, you're going to start looking at mediums that depend on how much it's raining or how much moisture is in the air and so on. You're going to have to make a decision on what your coding scheme is going to be or how you're going to communicate across here, depending on what that medium is. I think, as you mentioned Bob [Peile], that no one wants to be sitting there and trying to figure this out. The guy that's operating the system does not want to say, "Gee, I should select this because it's raining somewhat hard, or it's not raining at all or maybe there's some moisture in the air ...." Somehow you have to automate that. I see that as relatively slow adaptation, but something that is necessary to make the things work at the higher frequencies. That's been used or at least proposed in some satellite systems where you're talking down to a rainy ground station, you know, putting more and more coding on depending on what the medium is. I think that's an interesting area and where you can really affect slow adaptations, and it'll add a great deal to the flexibility of the systems.[2]

KEN WILSON: The comment on the customers' knowledge of the channel model I thought was ....

PEILE: I can't see who is talking .... Could you give me your name please?

WILSON: I had missed the point of Paul's presentation when he gave an extended presentation of the forest model. I thought that that was kind of ironic that we would say that immediately after he'd given us that model [LAUGHTER] and not only that, but

[2]H. Bustamante talked about this aspect in the last session.

nobody jumped up and down and said, "OK, now I can go out and build his equalizer for him to handle that model and enable his communications through that medium." So, you have a reply, sir ....

WELCH: Yes, I didn't really mean to make such a blanket statement; there are exceptions, some people know what their channel is. But there are plenty more that don't!

LINDSEY: Can I get back into the act? Sorry, I've got a technical comment. [LAUGHTER] It follows along with Gaylord's comments here. First of all, in the use of millimeter waves to go to higher frequencies I think that's certainly a proper direction to get on for many reasons we don't have to address here. The point is that coding buys you currently what is the number he gave over there. We know from 10^{-5} the best you can pick up is 10 or 11 dB, and right now maybe we're getting 6 or 7 dB of that.

When you have millimeter waves that are transmitted – in this case we're talking satellite communications again – we understand that when it rains you need to reserve in your power budget somewhere between 9 and 13 dB, depending on what frequency you're talking about, of course the level of the rain activity and what not, but it's usually 9-13 dB. Now, with regard to coming in with the system engineering solution to say that I'm going to play with coding above and beyond what I've already got there ... maybe I've missed your point. Then maybe you can comment and make my point.

HUTH: My point wasn't to get away from fading margins. You need rain margins. My point about rain, about the design of a system at this level, is that part of your gain – we've talked about this in terms of interception or somebody knowing you're out there and so on. A lot of the things that happen at millimeter waves is when you're trying to be covert and you want to put a rain margin in. You need to know how to control your power, your coding, the whole system, and there's not going to be somebody out there that's going to tell you. You know, if this isn't done automatically in some way, it's not going to be done. The way most people communicate, if given their shot, is to turn it up as high as they can get it and blast through. But if you're in a covert environment, and that's what you're trying to achieve, then you don't put a rain margin in – you don't do anything. You're trying to decide what your power control is, what your coding is, what is the signal design at that point – and that has to be automated somehow. I'm saying that adaptive coding is part of that game.

SASS: I just wanted to make a comment about Bill's comment about the lack of the reverse information flow from the customer. I can spend 24 hours a day trying to guide the contractors I'm involved with, but I think similarly each of you has bonds with the industrial community. The Army is buying its communication systems from probably four contractors: Hughes Aircraft, GD, ITT and GTE. Off the top of my head, these are the four primary contractors building all of these radio systems that I'll talk about tomorrow to some degree. You each have very close bonds with these contractors, and you ought to get to them and let them know your capabilities, and have them tell you what their problems are or what they envision as their next generation of problems. At the same time, I try to do the same thing.

PEILE: I think Paul's last point is a good subject for the Wrap-up Session, let's have an extended discussion of this thing. Are there any more questions on adaptive coding?

I would just like to mention something that

happened last week, as yet another application of adaptive coding. This came from the very high speed networking people. They made the point that when they start really whizzing along at gigabits, their biggest problem, because the switches are going to change their structure dramatically, might well be the missing packet problem. Instead of using ARQ, there was a suggestion that forward error correction in a network might be a better solution to the missing packet problem than the traditional solutions. It's funny how these arguments that start with jammers and interferers can come full circle and hit you where you're not expecting them.

I think we'll adjourn early.

<u>Wrap-Up Session</u>

Session Orgainzer: Robert Scholtz

Marvin Simon
New Ideas on Differential Detection with Multiple Symbol Memory

Paul Sass
Issues in Tactical Networks

Herman Bustamante
1105B Communication Subsystem for Satellite Channels

ROBERT SCHOLTZ: At a meeting of the Dean's Advisory Board at USC, Bob Lucky made a comment to me that it was surprising that communication theory has lasted this long. It's provided all of us guys with a great career but I'm surprised that we can still make money doing this stuff. But Bob Lucky considers himself out of that field for quite some time now. However Marvin [Simon] is still going great guns and he has an interesting talk for us this morning. Then I'll have Paul Sass talk more about where the army's plans are for communications in the future. And that may lead us up into the wrap up part of the session. I've asked each of the panel chairmen to put down a few thoughts about what they think basic research ought to be in their areas. Is there any left, or is it dead? What are the difficult problems that are remaining in some of these areas? Some will be easy to define. Some will be very difficult I think. Some of the areas are very mature that we've talked about. Some are just hidden, some need to be digested. And then, if anyone has any critiques or comments on the Workshop, I'd very much like to hear them.

Milly, Ms. Montenegro, do you have any comments you want to make to this group at this point administratively?

[RESPONSE NOT RECORDED]

SCHOLTZ: I'd like to say one thing with regard to Milly. This is the third workshop I've done and I've really come to rely a lot on her for all of the administrative work so I hope we'll all give her a hand right now. [LOUD APPLAUSE] Thank you very much. And Marvin ... take over.

MARVIN SIMON: *New Ideas on Differential Detection with Multiple Symbol Memory*

When Bob asked me for a title for this talk, I called it "New Ideas in Differentially Coherent Detection." Actually that might be a bad title. When I show you the idea that I would like to address today, I would venture to say that many of you would think that it is really an old idea and no doubt has been done before; interestingly enough, that was my reaction when I first thought about it too. But in looking through the literature, I really couldn't find where this was formally addressed although it's hinted at in various places. So let me show you what I've got and I'll let you people be the judge as to whether it's new or old.

Now, these comments are not intended to imply that the talk applies only to 2 PSK. Actually it applies to MPSK.

Let me start out by reviewing the notions of what I refer to as conventional differential detection and I'll show you where I go from there. With conventional differential detection, of course, you observe the received signal over two symbol intervals which I'll call the current and the previous, and you use the previous symbol as reference to demodulate the present symbol which we'll refer to as the current symbol. In order to do that, you must have differential encoding at the transmitter since you're looking at the difference of phases between adjacent symbols. From the standpoint of bit-error probability, if you look at MPSK and do what we call conventional differential detection, then you can derive asymptotically that the bit-error probability behaves according to this function. [VIEWGRAPH #1] This is a result that Dr. Pawula got a few years ago and represents a very accurate approximation. In the binary case the error probability turns out to be, as you all know, exponentially dependent on E_b/N_0, whereas if you do coherent detection in the binary case, you know the error

Multiple Symbol Differential Detection of MPSK

by

Marvin Simon
Dariush Divsalar

Jet Propulsion Laboratory

Multiple Symbol Differential Detection of MPSK

- Conventional Differential Detection of MPSK -

  - Observe the received signal plus noise over two symbol intervals (the current and the previous)

  - Use the *previous* symbol as a demodulation reference for the *current* symbol - *partially-coherent detection.*

  - Requires differential encoding of the input data phases.

  - Bit error probability performance, P_b, varies asymptotically as

$$P_b \cong \frac{1}{\log_2 M}\sqrt{\frac{1+\cos\frac{\pi}{M}}{2\cos\frac{\pi}{M}}}\,\mathrm{erfc}\sqrt{\frac{E_s}{N_0}\left(1-\cos\frac{\pi}{M}\right)};\; M \geq 3$$

where E_s = energy/symbol, N_0 = noise spectral density. For binary modulation (M = 2), $P_b = \frac{1}{2}\exp\left(-\frac{E_b}{N_0}\right)$.

VIEWGRAPH #1

probability goes like a complimentary error function of E_b/N_0. There's a fairly significant difference between those two behaviors. Let me just give you a few numerical points which serve as a degree of merit. [VIEWGRAPH #2]

If you look at the binary case, then at $P_b = 10^{-5}$ the difference between the two is about 3/4 dB. At 4 PSK the difference is about 2.2 dB, at 8 PSK the difference is about 2.5 dB, and as you go to larger and larger M it asymptotically approaches a 3 dB difference. What I want to talk about today is a means of doing differentially coherent detection but trying to essentially narrow down that performance gap, trying to approach coherent detection without actually having to track or acquire a carrier, or things like that. [VIEWGRAPH #3] The scheme really involves what's referred to as *multiple symbol differential detection* and I'll just summarize the ideas and then we'll go into the actual discussion. Basically what we do is we observe the received signal over N-intervals rather than 2 intervals, and instead of estimating symbol by symbol, we indeed use a maximum-likelihood sequence algorithm to detect the most current $N-1$ symbols. The interesting thing is that we still use the first symbol in the sequence as the reference. So we have N symbols, we use the first as a reference, and then we detect the next $N-1$ in a sequence estimation. The more interesting point I think, too, is the fact that the encoding for this scheme is still the exactly same differential encoding you would use as if you had conventional differential detection which, by the way, would correspond to $N = 2$, the 2-symbol observation. There's an interesting implication behind that which I'll mention later. [VIEWGRAPH #4]

So now what we're trying to do is bridge the gap between differential detection (2-symbol observation) and coherent detection. It turns out that we can actually fill that gap entirely and how much we fill that gap depends upon the length of the sequence we observe which I am denoting by n. You'll see that numerically later on. The interesting part, of course, and again this is not very surprising, is that you can formally show that as n goes to infinity, i.e., as you include more and more memory or as you do a longer sequence estimation, you indeed approach coherent detection exactly. How close you come to that limit is the issue because the scheme which I'm going to talk about has a complexity which grows, in effect, like MN. So in that sense you're trading complexity with performance but this is probably the most significant point, and that is, practically speaking with very few additional memories like 1 or 2, you can start to approach coherent detection. You'll see that again numerically, as I said, a bit later. [VIEWGRAPH #5]

So let's look at the mathematical model for the problem. We're going to start out by transmitting MPSK which is a constant power, constant envelope signal whose possible phase angles can be viewed as m points uniformly distributed around a circle. The received signal is the transmitted signal vector s_k times a phase angle $e^{j\theta_k}$ where θ_k represents the unknown phase of the channel. Again, in the absence of any information, we assume this to be uniformly distributed and of course we have the usual additive Gaussian noise. So the problem is, as is true in most maximum-likelihood sequence estimation algorithms, we observe a received of N symbols sequence which I'll call $\underline{r}$. In conventional differential detection, you have to assume that the unknown phase is constant over 2 symbols. Here you have to assume that

Comparison of Differentially Coherent and Coherent Detection of MPSK

- At $P_b = 10^{-5}$:

  - <u>M = 2</u>

    $(E_b/N_0)_{diff.} = (E_b/N_0)_{coh.} + 0.75$ dB

  - <u>M = 4</u>

    $(E_b/N_0)_{diff.} = (E_b/N_0)_{coh.} + 2.20$ dB

  - <u>M = 8</u>

    $(E_b/N_0)_{diff.} = (E_b/N_0)_{coh.} + 2.50$ dB

VIEWGRAPH #2

Multiple Symbol Differential Detection of MPSK

- Multiple Symbol Differential Detection of MPSK -

  - **Observe the received signal plus noise over N (more than two) symbol intervals.**

  - **Use a maximum-likelihood sequence estimation algorithm to detect the current N-1 symbols (the first symbol again acts as the reference phase)**

  - **Requires *identical* differential encoding of the input data phases as for conventional differential detection.**

VIEWGRAPH #3

Multiple Symbol Differential Detection of MPSK (cont'd)

- Error probability performance varies between that of conventional differential detection and that of coherent detection depending on the value of N.
  - Theoretically, in the limit of infinite symbol observation, the error probability performance becomes *identical* to that corresponding to coherent detection of MPSK with differentially encoded input phases.
  - Practically, with only a few additional observation intervals, one can approach coherent detection performance.

VIEWGRAPH #4

Mathematical Model for Maximum-Likelihood Detection of MPSK

- Transmitted signal in interval $kT \leq t \leq (k+1)T$

$$s_k = \sqrt{2P}e^{j\phi_k}$$

  - ϕ_k takes on one of M uniformly distributed values $\beta_m = 2\pi m/M$; $m = 0, 1, 2, \ldots, M-1$ around the unit circle.

- Received signal

$$r_k = s_k e^{j\theta_k} + n_k$$

  - n_k is a sample of zero mean complex Gaussian noise with variance $\sigma_n^2 = 2N_0/T$.
  - θ_k is an arbitrary phase introduced by the channel which is assumed to be uniformly distributed in $(-\pi,\pi)$.

- Consider a received sequence, $\underline{r}$, of length N symbols and assumed that θ_k is independent of k over the length of this sequence, i.e., $\theta_k = \theta$.

$$\underline{r} = \underline{s}e^{j\theta} + \underline{n}$$

VIEWGRAPH #5

the unknown phase is constant over N symbols. Now of course, the validity of this assumption relates to the dynamics of the channel. How long can you essentially assume the phase to be constant? And that essentially gives you an idea of how much memory you can use in this kind of a scheme. So based on this we actually want to make a maximum-likelihood sequence decision rule. [VIEWGRAPH #6] So what we do, as all of you know, we form the a posteriori probability of the received vector given the transmitted vector, and since we don't know the phase θ we average over that phase which happens to have a uniform distribution. This conditional probability distribution is Gaussian. If you go through some steps you can come up with a metric, or decision rule, which says, as you well know, to correlate the received set of signal samples with all the possible transmitted phases, take the magnitude squared of that sum, and make a decision according to the maximum.

What I'm going to do is show you an interesting interpretation of this and also show you where differential detection really fits into this thing as I see it. You obviously notice that because this is a metric which involves e^j of the transmitted phases, I can add any phase, or essentially I can rotate the constellation if you like by any phase angle, and I don't change the metric because a shift in each of these phases by a fixed amount doesn't change the squared magnitude of that sum. So let me choose this arbitrary phase that I'm going to shift by as the first phase in the sequence. In that case, equivalently I can look at this method which now says correlate the received signal samples with the *difference* in phase between the phase of each symbol and the very first phase in the sequence. You can see that pictorially, again, so far there is nothing new but perhaps the formulation. I'll tell you when something new takes place. [VIEWGRAPH #7] Really what I'm saying then, is that to compute that metric, what you want to do is compare each of the transmitted phases in the sequence with the very first phase, which arbitrarily I've selected as my reference, and then multiply e^j of those phase differences times each of the possible received signal samples plus noise. Now the point is, there's still a phase ambiguity because I can phase rotate by any angle and I still haven't changed anything. So how do we resolve the phase ambiguity? The way we do that is to use classical differential encoding at the transmitter. The information symbols, $\Delta\phi$'s, are then given as the difference between two adjacent transmitted phases. So, if this is true, then the phase difference between any phase in the sequence and the very first phase will be the sum of the pairwise adjacent phases which of course would just be the sum of the information symbols. So that metric that you saw before now becomes the correlation of the received signal plus noise samples with the sum of all the information symbols from that point back, because basically you're looking at the difference between that phase and the first which is the sum of all the information symbols up to that point. I can interpret this in an interesting way. Since e^j of a sum of phase angles can be written as the product of e^j of the phase angles, then the correlation metric that I'm computing is really the correlation between any particular received signal plus noise sample and the product of all of the transmitted symbols up until that point starting at the beginning of the sequence. What I'm really doing is to take any point in the sequence and take a particular correlation of the received signal plus noise sample

Maximum-Likelihood Sequence Decision Rule

- Choose $\underline{s}$ that maximizes the a posteriori probability

$$p(\underline{r}|\underline{s}) = \int_{-\pi}^{\pi} p(\underline{r}|\underline{s},\theta)p(\theta)d\theta$$

$$p(\underline{r}|\underline{s},\theta) = \frac{1}{(2\pi\sigma_n^2)^{N/2}} \exp\left\{-\frac{\left\|\underline{r}-\underline{s}e^{j\theta}\right\|^2}{2\sigma_n^2}\right\}$$

$$p(\theta) = \frac{1}{2\pi}; \quad -\pi \le \theta \le \pi$$

- Decision rule:

choose $\hat{\underline{\phi}}$ if $\left|\sum_{i=0}^{N-1} r_{k-i} e^{-j\hat{\phi}_{k-i}}\right|^2$ is maximum

where $\underline{\phi}$ is a particular sequence of the input phases.

Note : the decision rule has a phase ambiguity associated with it since an arbitrary phase rotation, say ϕ_a, of all N components of $\underline{\phi}$ results in the same decision for $\underline{\phi}$. Thus, letting $\phi_a = \phi_{k-N+1}$, a sufficient decision statistic is

$$\eta \stackrel{\Delta}{=} \left|\sum_{i=0}^{N-1} r_{k-i} e^{-j(\phi_{k-i}-\phi_{k-N+1})}\right|^2$$

VIEWGRAPH #6

Maximum-Likelihood Sequence Decision Rule (cont'd)

$$\overbrace{\overbrace{\overbrace{\underbrace{\phi_{k-N+1}}_{=\phi_a}, \phi_{k-N+2}}^{\times r_{k-N+2}}, \cdots\cdots, \phi_{k-1}}^{\times r_{k-1}}, \phi_k}^{\times r_k}$$

sequence of N phases

- Phase Ambiguity Resolution - Differentially Encode Input Data

$$\phi_k = \phi_{k-1} + \Delta\phi_k$$

- Decision Statistic

$$\eta = \left| r_{k-N+1} + \sum_{i=0}^{N-2} r_{k-i} e^{-j\sum_{m=0}^{N-i-2}\Delta\phi_{k-i-m}} \right|^2$$

$$= \left| r_{k-N+1} + \sum_{i=0}^{N-2} r_{k-i} \prod_{m=0}^{N-i-2} e^{-j\Delta\phi_{k-i-m}} \right|^2$$

VIEWGRAPH #7

at this point times the product of e^j of all of those previous signals. We'll see why that comes about. So, this is indeed the metric that you would compute. To make a decision, you would take that sequence of $\Delta\phi$'s which maximize this metric. [VIEWGRAPH #9]

Let me show you some special cases of this to apply your thinking. Let's look at what happens if we just have one term. If we just take one term in that sum, we get a metric that's independent entirely on phase. There's no phase information in this, and thus this is classical noncoherent detection. If we look at $N = 2$ that is the present symbol, and the previous symbol times e^j of the present symbol, we actually get 3 terms. The first two terms are independent of the phase information, and thus we make a decision based on the third term. This is *classical differential coherent detection.* So in that context, you can think of conventional differential detection as being maximum-likelihood sequence estimation where the sequence is actually two symbols long, with unknown phase and you get indeed the conventional differential detection algorithm which you all know looks like this. [VIEWGRAPHS #10,11] In other words, you take the received signal plus noise, multiply it by the complex conjugate of the previous symbol plus noise, compare it to all the possible transmitted phase points around the circle and choose that phase which gives the largest real part. So again, there's nothing new here, only perhaps the formulation or the context in which I'm talking about it. [VIEWGRAPH #12]

Now the interesting part happens when you look at $N = 3$. If I look at 3 terms in the metric, again I have 2 symbols back, one symbol back weighted by the phase of the previous symbol, and the present symbol weighted by the sum of the present symbol and the previous symbol. Again you've got to remember, I said that we add up all of the phase prior to the point that we're observing in the sequence. Now if you do a little algebra on that, you come down to the following decision rule which chooses a pair of phases (you're observing 3 symbols and you're going to decide on the most recent two). So choose the present symbol and the previous symbol jointly according to this algorithm. Now let's look at those 3 terms because this is kind of interesting. The first term represents classical differential detection, the present symbol and the previous symbol. The second term is also classical differential detection of the previous symbol and the one before that. These two terms are what you'll do if you observe $\Delta\phi_k$ and $\Delta\phi_{k-1}$ alone and perform symbol-by-symbol detection. But the fact that we're doing a joint decision on $\Delta\phi_{k-1}$ and $\Delta\phi_k$ means we actually have to look at another term. And notice what that terms is. The term is essentially a hybrid, if you like, of the first two terms. It's the present symbol correlated with the symbol two symbols back, weighted by the sum of these two phase angles upon which I'm trying to make a decision. So that third term is the extra term you need because you're doing a sequence estimation as opposed to symbol by symbol detection. So let's look at the structure that this leads to. [VIEWGRAPH #13] Again I'm not saying that this is the structure you build, but this is the structure that is easiest to explain as to what is really going on. Notice that what's in the dashed box is essentially the first term in that expression. In other words this is classical differential detection of just the present phase based on a two-symbol observation. The dotted box represents a classical 2-symbol observation for the previous

Leib and Pasupathy IEEE Transactions on Information Theory, November 1988 (Distributed May 1989)

- Symbol by Symbol Detection as opposed to Maximum-Likelihood Sequence Estimation

  - Decision on $\Delta\phi_k$ (maximizing η over the M possible values of $\Delta\phi_k$) requires knowledge of previous N-2 transmitted symbols $\Delta\phi_{k-1}, \Delta\phi_{k-2}, \ldots, \Delta\phi_{k-N+2}$.

  - Used previous decisions $\Delta\hat{\phi}_{k-1}, \Delta\hat{\phi}_{k-2}, \ldots, \Delta\hat{\phi}_{k-N+2}$ for $\Delta\phi_{k-1}, \Delta\phi_{k-2}, \ldots, \Delta\phi_{k-N+2}$ in η - *decision feedback scheme* - subject to error propagation

- Computed error probability performance assuming *no* errors in previous decisions, i.e., perfect knowledge of the previous transmitted symbols - very optimistic.

VIEWGRAPH #8

Special Cases

- N = 1

$$\eta = |r_k|^2$$

Independent of the input data phases and thus cannot be used for making decisions on differentially encoded MPSK modulation - well-known case of *noncoherent* detection.

- N = 2

$$\eta = \left|r_{k-1} + r_k e^{-j\Delta\phi_k}\right|^2 = |r_{k-1}|^2 + |r_k|^2 + 2\,\mathrm{Re}\left\{r_k r_{k-1}^* e^{-j\Delta\phi_k}\right\}$$ 410

  - Decision Rule (conventional differential detection)

$$\text{choose } \Delta\hat{\phi}_k \text{ if } \mathrm{Re}\left\{r_k r_{k-1}^* e^{-j\Delta\hat{\phi}_k}\right\} \text{ is maximum}$$

Optimum Receiver in the sense of minimizing symbol error probability given that the unknown carrier phase is constant over two symbol times.

VIEWGRAPH #9

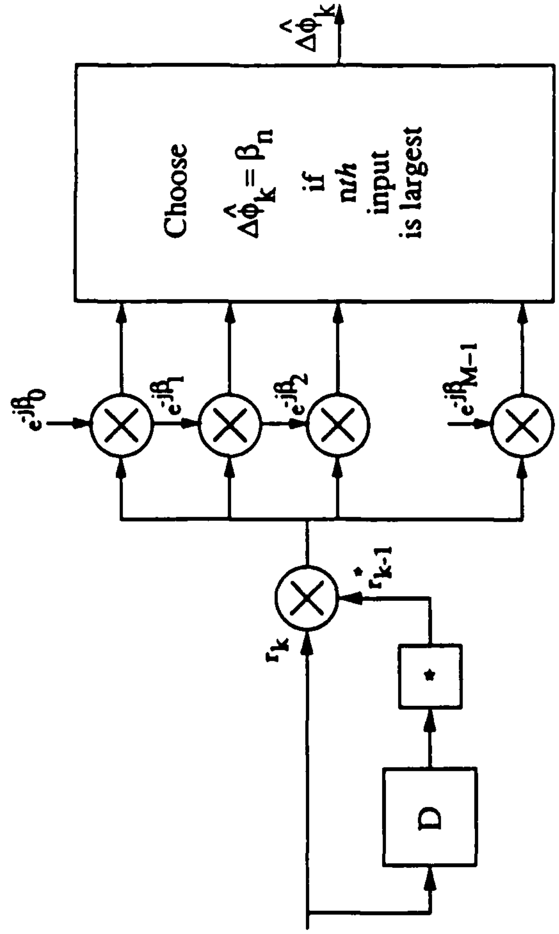


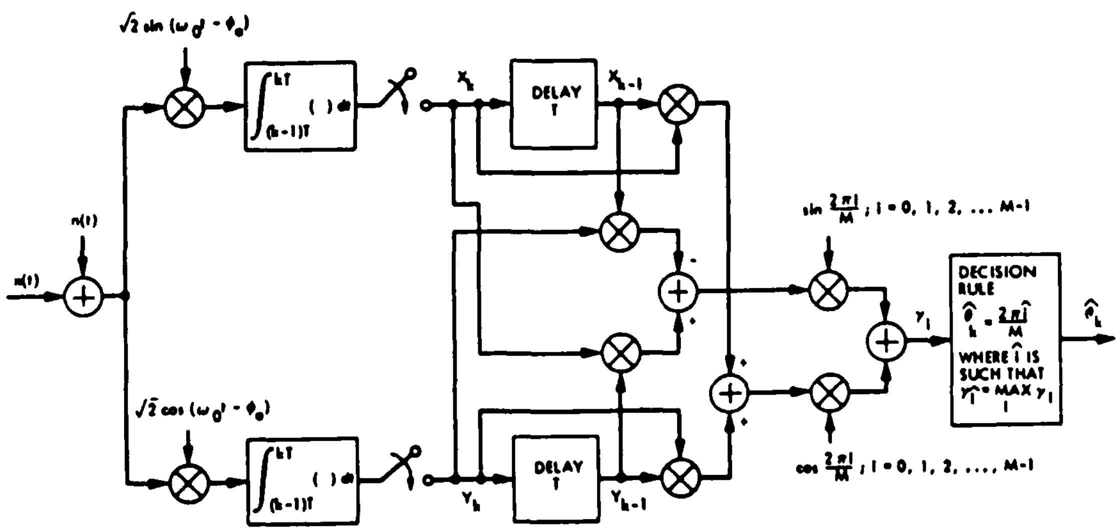
Figure 1. Conventional Differential Detector for MPSK

VIEWGRAPH #10

I-Q Implementation of DMPSK Receiver

$X_k = \mathrm{Re}\{r_k\};\ Y_k = \mathrm{Im}\{r_k\}$

$$\mathrm{Re}\{r_k r_{k-1}^* e^{-j\beta_i}\} = (X_k X_{k-1} + Y_k Y_{k-1})\cos\beta_i + (X_k Y_{k-1} - Y_k X_{k-1})\sin\beta_i$$
$$= \mathrm{Re}\{\xi_i\}$$



VIEWGRAPH #11

Special Cases (cont'd)

- N = 3

$$\eta = \left| r_{k-2} + r_{k-1} e^{-j\Delta\phi_{k-1}} + r_k e^{-j(\Delta\phi_k + \Delta\phi_{k-1})} \right|^2 =$$

$$\left|r_{k-2}\right|^2 + \left|r_{k-1}\right|^2 + \left|r_k\right|^2 + 2\,\mathrm{Re}\left\{ r_k r_{k-2}^* e^{-j(\Delta\phi_k + \Delta\phi_{k-1})} \right\}$$

$$+ 2\,\mathrm{Re}\left\{ r_{k-1} r_{k-2}^* e^{-j\Delta\phi_{k-1}} \right\} + 2\,\mathrm{Re}\left\{ r_{k-1} r_{k-1}^* e^{-j\Delta\phi_k} \right\}$$

- Decision Rule

choose $\Delta\hat{\phi}_k$ and $\Delta\hat{\phi}_{k-1}$ if

$$\mathrm{Re}\left\{ r_k r_{k-1}^* e^{-j\Delta\hat{\phi}_k} + r_{k-1} r_{k-2}^* e^{-j\Delta\hat{\phi}_{k-1}} + r_k r_{k-2}^* e^{-j(\Delta\hat{\phi}_k + \Delta\hat{\phi}_{k-1})} \right\} \text{ is maximum}$$

First and second terms are identical to those used to make successive and independent decisions on $\Delta\phi_k$ and $\Delta\phi_{k-1}$. The third term is required for optimum *joint* decision on $\Delta\phi_k$ and $\Delta\phi_{k-1}$.

VIEWGRAPH #12

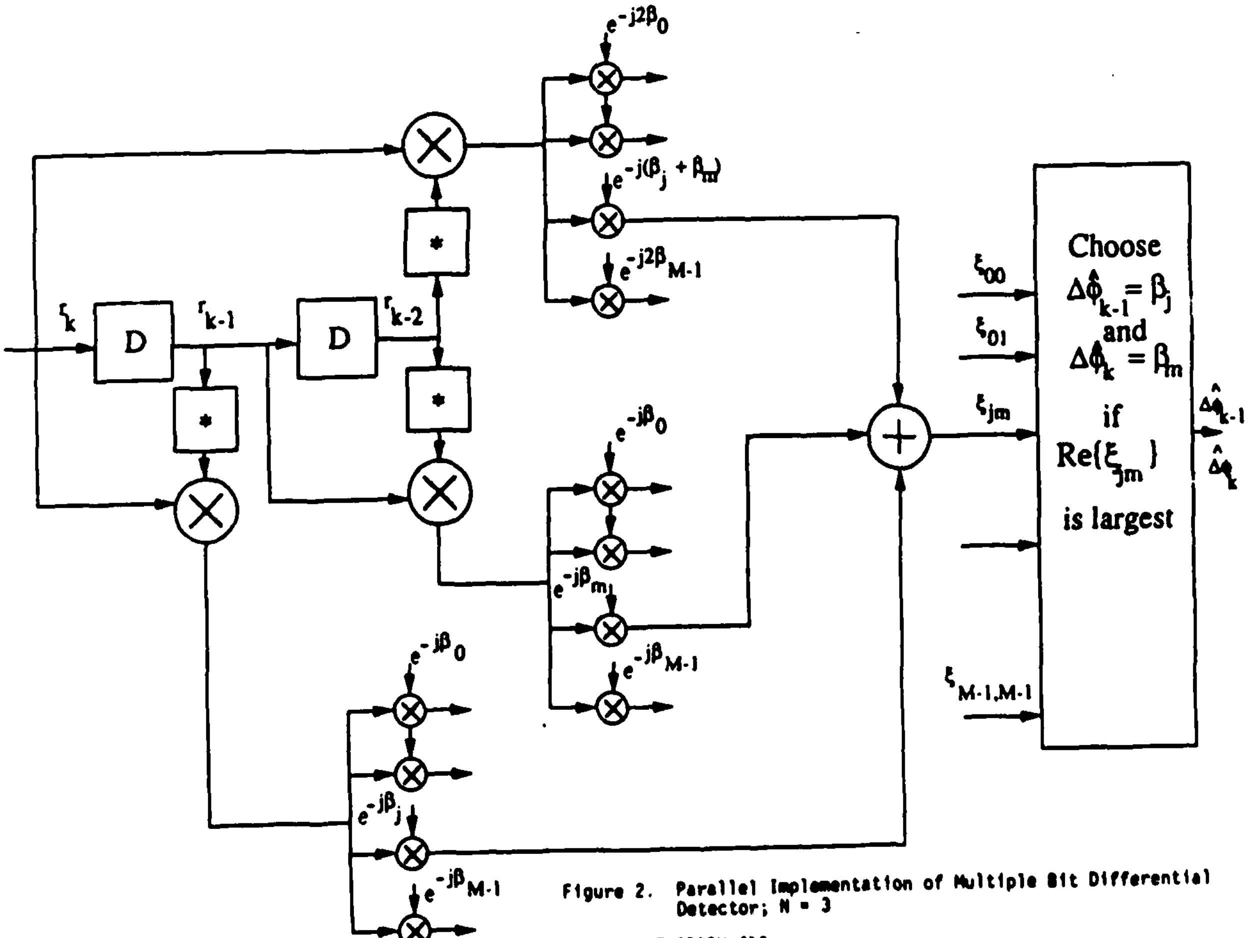


Figure 2. Parallel Implementation of Multiple Bit Differential Detector; N = 3

VIEWGRAPH #13

phase ϕ_{k-1}.

Now you've also got a third term here. We add these 3 terms taking ito account all possible combinations of the transmitted phases and we choose that pair of symbols as a sequence estimate which leads to the largest real part of that sum. So this is essentially the structure that you get based on one additional symbol observation. Again, I don't suggest that you build it that way. This is what we call parallel structure, but it does show the various components of this metric. I have a serial implementation which I'll just put up for a second. [VIEWGRAPH #14] It's quite a bit simpler. In the serial implementation, the dashed and dotted lines represent the same item as in the parallel implementation. And this essentially represents the new term that I have to add. You may wonder why the division by the magnitude of $|r_{k-1}|^2$. I can show you very briefly why that comes about. If you look at the third term in the metric, it represents the product of the first two terms, but I also pick up an additional $(r_{k-1})(r_{k-1})*$, that is, the magnitude squared of the middle symbol. So in order to get the third term, I can multiply the first two and divide by this middle symbol magnitude squared.

Here it is. It's the product of these two terms divided by the magnitude of the middle symbol-squared. Again, this is just another implemention. There are other ways. O.K., how does this thing perform. [VIEWGRAPH #15] That's perhaps the most interesting question to ask. How does it perform relative to conventional differentially coherent detection? When we're doing a sequence estimation, what we try to do is find an upper bound on the performance. We use a union bound here, and compare some particular transmitted sequence with some error sequence. We find the so-called pairwise error probability of the bit streams which generated those sequences. We weight that by the Hamming distance between the two bit sequences which generate the two sequences $\Delta\phi$ and $\Delta\hat{\phi}$ and then we divide that by the possible number of combinations. So the key element in evaluating the bit-error rate is evaluating this pairwise probability which is just a sequence pairwise probability based on the metric that I showed you. Now there's two approaches to doing this. One is trying to get an exact expression which in this particular problem you can, or you can upper-bound it by a Chernoff bound which is unfortunately the way we have to go when we do the coded case. (How am I doing on time? ... Less than half?) [VIEWGRAPH #16]

Let me just show you the result that we get (I don't want to bore you with equations but there's an interesting issue here) .... The pairwise error probability, and this is the result that Seymour Stein and many other people got years ago, can be expressed in terms of the Marcum Q-function of two arguments. Now this is the key parameter: E/N_0. The two arguments of the Q-function pertain to E/N_0 and the length of the sequence, and a parameter which is called δ. This δ seems to crop up in all the analyses of this kind of problem, and what it represents again, is that for each position in the sequence, take the product of all of the past correct symbols prior to that position and the product of all of the complex conjugates of the past error symbols prior to that past position, and multiply those two together. Now this is a nonlinear parameter and so the point I want to mention right here before I get to the coded case is when you apply this in a coded environment like trellis coding, Euclidean distance is not an appropriate measure. And

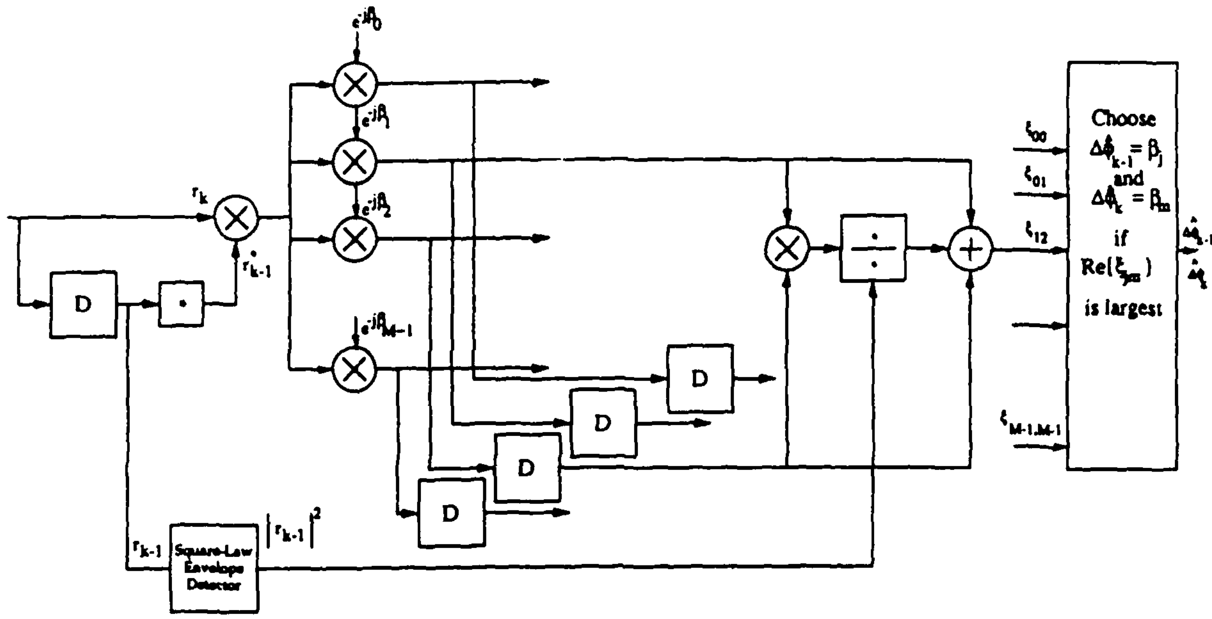


Figure 3. Serial Implementation of Multiple Bit Differential Detector; N = 3

VIEWGRAPH #14

Bit Error Probability Performance

Let u be the sequence of b = (N-1)$\log_2 M$ information bits that produces $\underline{\Delta\phi} = (\Delta\phi_k, \Delta\phi_{k-1}, \ldots, \Delta\phi_{k-N+2})$ at the transmitter and $\hat{\underline{u}}$ the sequence of b bits that result from the detection of $\underline{\Delta\hat{\phi}}$.

$$P_b \le \frac{1}{b}\frac{1}{M^{N-1}} \sum_{\underline{\Delta\phi}} \sum_{\neq \underline{\Delta\hat{\phi}}} w(\underline{u},\hat{\underline{u}}) \Pr\left\{\hat{\eta} > \eta | \underline{\Delta\phi}\right\}$$

- $w(\underline{u},\hat{\underline{u}})$ = Hamming distance between $\underline{u}$ and $\hat{\underline{u}}$.
- $\Pr\left\{\hat{\eta} > \eta | \underline{\Delta\phi}\right\}$ = pair-wise probability of error that $\underline{\Delta\hat{\phi}}$ is incorrectly chosen when indeed $\underline{\Delta\phi}$ was sent.

VIEWGRAPH #15

Pair-wise Error Probability
(Exact Evaluation)

$$\Pr\left\{\hat{\eta} > \eta \mid \underline{\Delta\phi}\right\} = \frac{1}{2}\left[1 - Q(\sqrt{b},\sqrt{a}) + Q(\sqrt{a},\sqrt{b})\right]$$

- Q(α,β) = Marcum's Q-function

$$\begin{Bmatrix} b \\ a \end{Bmatrix} = \frac{E_s}{2N_0}\left[N \pm \sqrt{N^2 - |\delta|^2}\right]$$

$$\delta \stackrel{\Delta}{=} \sum_{i=0}^{N-1} e^{j\sum_{m=0}^{N-i-2}\left(\Delta\phi_{k-i-m} - \Delta\hat{\phi}_{k-i-m}\right)} = \sum_{i=0}^{N-1}\left(\prod_{m=0}^{N-i-2} e^{j\Delta\phi_{k-i-m}}\right)\left(\prod_{m=0}^{N-i-2} e^{-j\Delta\hat{\phi}_{k-i-m}}\right)$$

VIEWGRAPH #16

Asymptotic Approximations to
Marcum's Q-Function

$$Q(\alpha,\beta) \cong 1 - \frac{1}{\alpha - \beta}\sqrt{\frac{\beta}{2\pi\alpha}}\exp\left\{-\frac{(\alpha-\beta)^2}{2}\right\}; \quad \alpha >> \beta >> 1$$

$$Q(\alpha,\beta) \cong \frac{1}{\beta - \alpha}\sqrt{\frac{\beta}{2\pi\alpha}}\exp\left\{-\frac{(\beta-\alpha)^2}{2}\right\}; \quad \beta >> \alpha >> 1$$

VIEWGRAPH #17

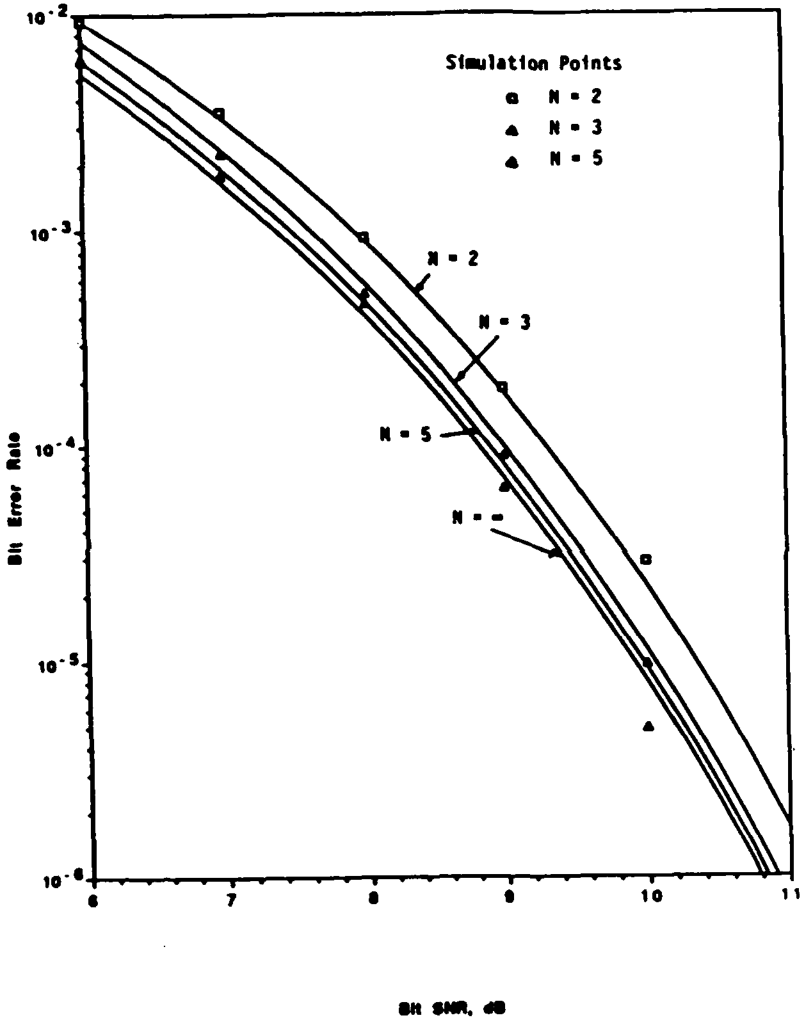
hopefully we'll have time to talk about what the right measure is. Now, if I put this back into the error probability expression, you'll get an answer in terms of Q-functions which doesn't give you a lot of insight into how this thing performs. So what we did is we used some well-known asymptotic approximations [VIEWGRAPH #17] to the Marcum Q-function, valid for large E/N_0, and lo and behold we get some very nice simple results. Let me show you. [VIEWGRAPH #18]

Let's look first at the case of the 3-symbol observation which is one additional memory. Let's look at binary PSK. Now you recall for binary DPSK, the bit-error probability is exactly an exponential of E/N_0. So I have factored that term out over here. This is classical conventional differential detection, and the term in front of this factor essentially reflects the *improvement in performance* by going to the 3-symbol observation. It's a very simple term. The most important thing to notice is that it decreases like $(\sqrt{E/N_0})^{-1}$. In other words, you get an additional error performance improvement with increasing signal-to-noise ratio. Note that you don't get any improvement in the error exponent. Now if you generalize that to arbitrary length sequences, instead of $N = 3$, you get again almost the same result. Again, it's the classical differential detection improved by this term which is very simple to gain some intuition from. Now look what happens for large N. For large sequence length, the term $\sqrt{(N-1)/(N-2)}$ goes to 1 as you can see, and all I have left, if I look at the asymptotic approximation for the error function, is coherent detection behavior. So that says that as E/N_0 approaches infinity, I indeed approach coherent detection performance. Now you may ask where's the factor of 1/2 in front of here that you're all used to seeing. Does anybody know why that's not there? Remember we are talking about differential encoding. So really the fair comparison is between this scheme and a differentially encoded PSK which as a factor of 1/2 when compared with non-differentially encoded PSK performance.

Let me just show you some curves on how this thing looks. [VIEWGRAPH #19] These are done using the exact expression, not using the asymptotic approximation. [QUESTION FROM THE BACK ... UNINTERPRETABLE] Because it turns out $N = 2$ is the singular case. It just doesn't fit into that expression. You're right, that expression is only valid if N is greater than 2. For $N = 2$ we know the exact answer. In short, when you approximate the Marcum Q-function the approximations that you have to make don't hold when $N = 2$.

So let's look at a few performance curves. Here's classical DPSK with symbol observation $N = 2$. Here's $N = \infty$ which is coherent PSK, and the interesting thing is if you just go to $N = 3$, one additional symbol, you pick up about 0.5 dB of the 0.75 dB difference at 10^{-5}. And if you go up to $N = 5$ you're pretty close to coherent performance. So the thing that I want to leave you with in so far as DPSK is concerned is that with very little additional memory, you can start to approach coherent performance very quickly. You buy most of the gain with the first additional symbol of memory. I won't show you the equations but you can generalize this to an arbitrary number of points on the circle. [VIEWGRAPH #20] Let me just show you some other curves. [VIEWGRAPHS #21,22] For $N = 4$, again here is coherent detection and here's classical DPSK. As I said, there's about 2.2 dB difference between them at 10^{-5}. If I go to one additional

Figure 3. Bit Error Probability versus E_b/N_0 for Multiple Differential Detection of MPSK; M = 2.



VIEWGRAPH #19

Bit Error Probability Performance (Special Cases)

- M = 2, N = 3 (Binary DPSK with three symbol observation)

$$P_b \lesssim \frac{2\sqrt{2}}{\sqrt{\pi \frac{E_s}{N_0}}}\left[\frac{1}{2}\exp\left\{-\frac{E_s}{N_0}\right\}\right]$$

- The factor in front of the bracketed term represents the improvement obtained by increasing the memory of the decision by one symbol interval from N = 2 to N = 3.

- M = 2, N arbitrary (Binary DPSK with N symbol observation)

$$P_b \lesssim \frac{2}{\sqrt{\pi \frac{E_s}{N_0}}}\left(\sqrt{\frac{N-1}{N-2}}\right)\left[\frac{1}{2}\exp\left\{-\frac{E_s}{N_0}\right\}\right]$$

- Asymptotically for N large

$$P_b \lesssim \frac{1}{\sqrt{\pi \frac{E_s}{N_0}}}\exp\left\{-\frac{E_s}{N_0}\right\} \cong \operatorname{erfc}\sqrt{\frac{E_s}{N_0}}$$

Same as asymptotic P_b of coherent detection of BPSK with differentially encoded input symbols

VIEWGRAPH #18

Bit Error Probability Performance (Special Cases)

- M arbitrary, N arbitrary

$$P_b \le \frac{2}{(\log_2 M)\sqrt{\pi \frac{E_s}{N_0}}} \left(\sqrt{\frac{N + |\delta|_{\max}}{|\delta|_{\max}(N - |\delta|_{\max})}} \right) \times \exp\left\{ -\frac{E_s}{2N_0}(N - |\delta|_{\max}) \right\}$$

$$|\delta|_{\max} = \sqrt{(N-1)^2 + 2(N-1)(1 - 2\sin^2\frac{\pi}{M}) + 1}$$

VIEWGRAPH #20

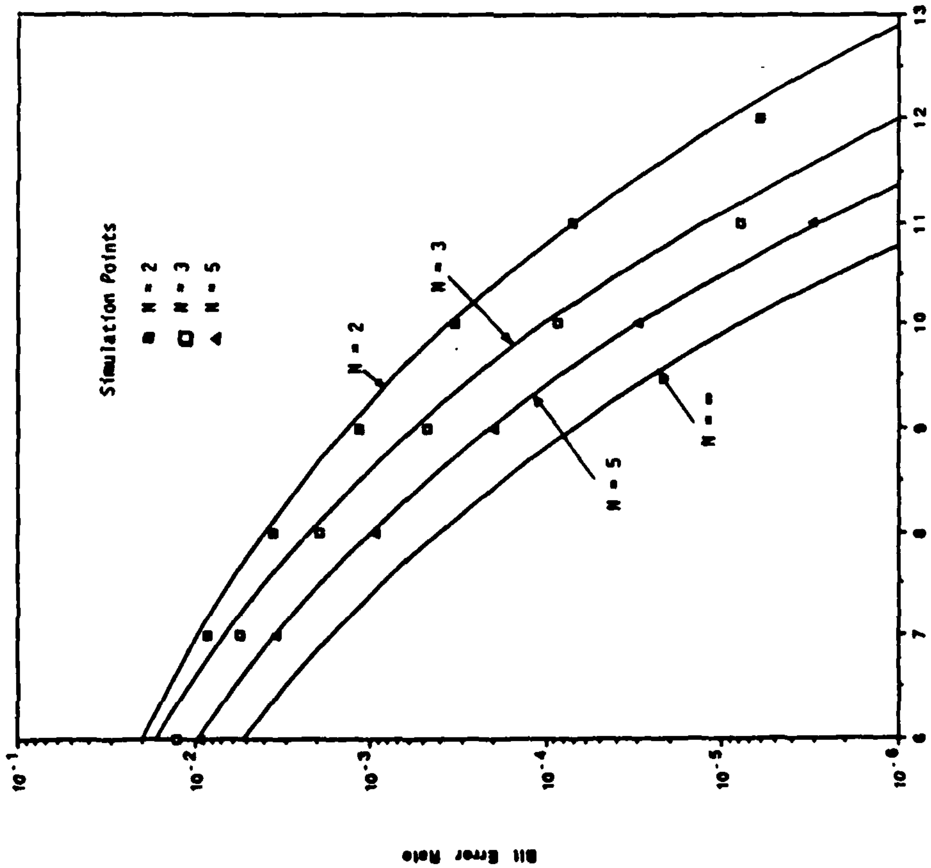


Figure 4. Bit Error Probability versus E_b/N_0 for Multiple Differential Detection of MPSK; M = 4

Bit SNR, dB

VIEWGRAPH #21

Applications of Multiple Symbol Differential Detection of MPSK

- Channels requiring Error Correction Coding to achieve necessary performance - results obtained for uncoded modulation have been extended to trellis coded modulations.

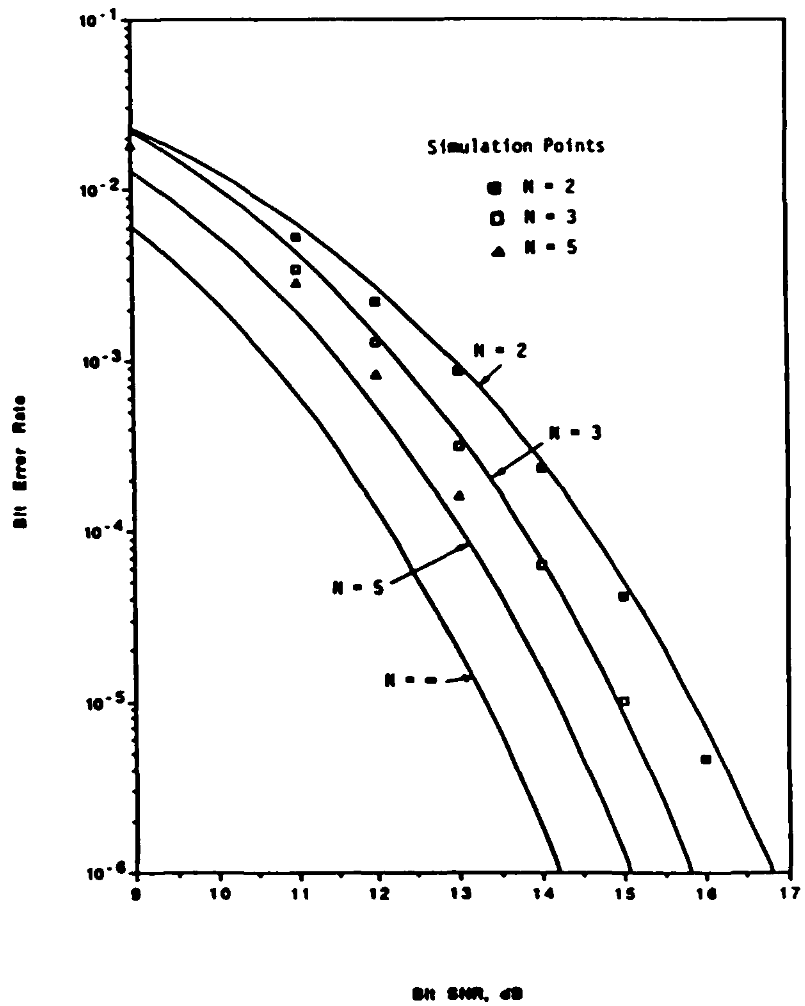
- Multipath Fading Channels - acquiring and tracking a coherent demodulation reference signal is difficult if not impossible. Fading must however be slow enough so that the phase of the received carrier can be assumed constant over the number of symbols corresponding to the observation.

- Satellite Sound Broadcast Channels - both time and frequency selectivity are issues. Use of trellis coding combined with multiple symbol differential detection of MPSK would allow for more channels to be accommodated than with presently used techniques (i.e., conventional differential detection of convolutionally encoded MPSK)

- Benign Channels - e.g., AWGN. Multiple symbol differential detection leads to a considerably simpler implementation than coherent detection (no carrier acquisition and tracking loop required) but with much less performance penalty than conventional differential detection.

VIEWGRAPH #23

Figure 5. Bit Error Probability versus E_b/N_0 for Multiple Differential Detection of MPSK; $M = 8$.



VIEWGRAPH #22

symbol memory, I pick up about 1 dB of that, and as I said for $N = 5$ you start to approach coherence. By the way, the points on there are simulation results and they fit very nicely with these tight upper bounds that I got by approximating the Marcum Q-functions. As you can see, they are very, very close. [VIEWGRAPH #24]

Let me just briefly, in 5 minutes tell you how this can be applied in a trellis-coded environment which is really what our motivation was. (Maybe I don't need the pointer.) I won't give you the details. In a trellis-coded environment, a few years ago we introduced something called a multiple trellis-code. A multiple trellis code (maybe, I should tell you that first) is a code where instead of outputing 1 M-ary symbol at a time, essentially you encode a set of input bits b into a set of s output symbols. You organize these s symbols into k groups of $\log_2 M$ symbols each, and you output k M-ary symbols at a time. So the throughput is b/k which is a rational rather than an integer throughput as in the conventional Ungerboeck-type code. This gives you a lot of flexibility in designing a trellis code. The point is, that in this kind of a trellis code, instead of having one symbol assigned to every trellis branch, you have a sequence of symbols (k) assigned to each trellis branch. As you can see, it lends itself nicely to the multiple symbol DPSK detection scheme. What I actually do is construct such a multiple trellis code and assign the so-called multiplicity, or number of symbols per trellis branch equal to $N - 1$, the number of symbols that I am trying to estimate with my multiple symbol algorithm. So in a sense, you can look at the trellis coding application as a combination of a sequence or sub-sequence assigned on a trellis code on which I apply the uncoded metric that I had before. In other words I take this uncoded metric, I apply that to each branch of the trellis, and then I go ahead and figure out what the appropriate distance structure is which, unfortunately, I don't have time to talk about. Then you go ahead, and this here is the interesting part. I'll just mention it without the details. It turns out that even though the distance metric is not squared Euclidean distance, I can construct, (and this is just a mathematical construction, this is not actually the code you would build) another trellis code with a slightly larger multiplicity for which the Euclidean distance is again the proper metric. Therefore I can use all the analysis techniques that I used before. I can go ahead and determine the error probability based upon well-known techniques. So I think the key point is the application to a multiple trellis code and the fact that you can indeed construct an equivalent trellis code which has Euclidean distance measure.

I'll just sum up with one comment which I would like to leave you with. [VIEWGRAPH #23] Most of you know many applications where differential detection is useful. The one that I'd like to at least suggest to you as food for thought is on the benign channel where we typically would use coherent detection requiring carrier acquisition, carrier phase tracking, lock detection, and other attendant functions. The reason people have shied away from differentially coherent detection systems like this was things like two and a half dB penalty when you're using 8PSK. What I would like to suggest possibly is using a scheme like this which is relatively simple except for the additional complexity introduced by the maximum-likelihood algorithm, that is, using differentially coherent detection in a system where you might think of using only coherent detection. And with that I'll

Application to Trellis Coded Systems

- Construct a *multiple* trellis code with multiplicity k = N-1, i.e., N-1 MPSK symbols are assigned to each branch of the trellis diagram.

- Associate the uncoded decision metric η with each branch of the trellis.

- Determine the appropriate distance measure (no longer squared Euclidean distance as for coherent detection in AWGN).

- Construct an equivalent multiple trellis code (with larger multiplicity) for which squared Euclidean distance is the correct distance measure.

- Determine the bit error probability performance of this equivalent trellis code by previous analysis techniques for trellis coded conventional differentially detected modulations.

VIEWGRAPH #24

stop.

SCHOLTZ: Very good, Marv. Now it occurs to me that this talk has been very different from other talks that we've heard so far, and that's why it's given this morning. Now Marv, when you and I talked on the phone, it seems to me you said it would be interesting if someone else here could point out where this had been done before. And, for example, I was thinking about some of the stuff we heard on the equalization panel where they talked about despinning signals. I was wondering if there was any relation between this despinning process, where you're taking out carrier rotations on successive samples, that is closely related to what Marv is talking about.

[SOMEONE IS COMMENTING WHICH IS NOT BEING RECORDED.]

WILLIAM LINDSEY: I saw John's hand up. I have two comments, one relative to your statement there. Marvin's discussion here I think that the notion of rotation of the symbols as a consequence of frequency offset can be included but it isn't included in his model at the moment. He's assuming frequencies are known so that the symbols don't rotate when they come in. They are just offset by some constant phase over the n symbols. It strikes me that he could perhaps put the frequency offset in. In which case there would be a bias in some things happening. He can analyze that I'm sure. I'd like to make a statement relative to Gauss .... [LAUGHTER] When you started out you mentioned it was or wasn't new. I think that this idea, where it probably has been looked at but not to the extent that you've been motivated to go into the direction of it in, in particular, it is my belief that in 1962, Paul Wentz and John Hancock took a look at the notion of using multiple symbols to improve differentially coherent detection. Where that was published I believe is not in a journal, but in a conference, and I have to look it up if you're interested in looking at that. And more recently, in S.L. Rice's memorial proceedings on Information Theory, there's a paper where he in essence is attempting to exploit this concept in MSK environment using differential encoding. I have to think, based on having read that paper and having seen your work for the first time here that your analysis is very clean and there should be some coupling and abridging between those two pieces of work.

SIMON: Let me comment if I can on Bill's comments on the paper which just came out. There's a paper by Leib and Pasupathy which is in the November issue which just came out. I just got it about two days before I came out here so I did prepare this viewgraph [VIEWGRAPH #8] and I'm glad you did bring it up. They essentially come up with the same metric interestingly enough but the way they handle it is they don't do maximum-likelihood sequence estimation. They essentially assume they have knowledge of all the previous symbols, and indeed do what I would call a *decision-feedback* kind of scheme. They are still doing symbol-by-symbol detection based upon previous knowledge of all the previous symbols and in their analysis, unfortunately, they do assume they have perfect knowledge of all previous symbols. They don't allow for errors in the previous symbols. So their results are very optimistic. But indeed that's another twist to this whole scheme. Instead of doing a sequence estimation, as we've heard before, you can actually do a decision feedback type of algorithm.

JOHN PROAKIS: O.K. Marv, I think your comment here with regard to the use of decision feedback is a good point. I was wondering if you had made any comparisons

Leib and Pasupathy IEEE Transactions on Information Theory, November 1988 (Distributed May 1989)

- Symbol by Symbol Detection as opposed to Maximum-Likelihood Sequence Estimation

  - Decision on $\Delta\phi_k$ (maximizing η over the M possible values of $\Delta\phi_k$) requires knowledge of previous N-2 transmitted symbols $\Delta\phi_{k-1}, \Delta\phi_{k-2}, \ldots, \Delta\phi_{k-N+2}$.

  - Used previous decisions $\Delta\hat{\phi}_{k-1}, \Delta\hat{\phi}_{k-2}, \ldots, \Delta\hat{\phi}_{k-N+2}$ for $\Delta\phi_{k-1}, \Delta\phi_{k-2}, \ldots, \Delta\phi_{k-N+2}$ in η - *decision feedback scheme* - subject to error propagation

- Computed error probability performance assuming *no* errors in previous decisions, i.e., perfect knowledge of the previous transmitted symbols - very optimistic.

VIEWGRAPH #8

with a decision feedback scheme?

SIMON: No, because I just saw this.

PROAKIS: I believe that's an old scheme. It goes back to some work done by Bob Price in the early 60s and perhaps Bill Lindsey may have looked at this problem in a number of papers that he published around 62-64. That's the conventional scheme. I think the degradation in performance due to decision errors is really very minor and that you begin to see large gains in performance just by looking at phase estimation over a very small number of symbols.

SIMON: I think as all decision feedback is concerned, that's a function of what error rates you are looking at. In a coded system where you might be looking at 10^{-3} symbol error rate, I think decision error rates are very important. And I think my scheme will win out in that environment. So again, I don't want you to think only in terms of the uncoded environment, you have to think of the coded phase case too.

SEYMOUR STEIN: Marv, if you move the decision to the middle of the sequence, so that you are looking at a total sequence and trying to make a decision about the symbol in the middle, I think there is a very close relationship to some work using similar kinds of estimation for phase coherent FSK extending over multiple symbols.

SIMON: Yes, but again, the difference in CPM is you have built-in phase continuity because of the way you generate CPM. Here the phase continuity is built in because of the channel phase being kept constant. That's what introduces the correlation between symbols. And there is a close relation conceptually, but I don't think anyone has formally, that I have seen, addressed what the gains are in differential detection. So I agree with you. Conceptually there's a very close relationship. That's why I say it's been mentioned and pointed out in a lot of places but I haven't seen it formally addressed.

JOHN CIOFFI: Maybe I misunderstood, but on the decoder slide that you showed, you said the complexity went up as m^n. And I would think you could describe this with an m-state trellis and basically ....

SIMON: No, I don't think you can because its a non-linear algorithm. I don't think you can decode this with an m-state trellis.

CIOFFI: You have impossible previous phases that you could be in, which would be the error states, and then you just determine the metric on the basis of each of those possible previous states rather than working out what the metric is for each of the sequences is independently.

SIMON: Well maybe you can. I really haven't thought about that.

GAYLORD HUTH: I have a question. I am interested in your scheme for a frequency hopping system. In this case, you are trying to estimate the phase over the hop. In the past, we have investigated ordinary differential demodulation and forward coherent estimation. You cannot use a loop because of the short duration of the hop. If I am going to use all the symbols (or a large number of them) in the hop, why would I want to use differential estimation when I can use coherent forward estimation across the hop? Are you saying I am better off complexity-wise?

SIMON: Yes, again as in any coherent scheme, only in the sense that this is only intended for schemes where you need fast carrier acquisition and where you perhaps don't want the complexity of having to do acquisition, re-acquisition, tracking, whatever.

HUTH: What I was thinking is that in a frequency hop system, I am looking at a single hop at a time and I am performing

a phase estimation over the symbols in that hop. I am not concerned about more than one hop at a time because the hops have independent phases. Assume that the phase does not change over that hop. Now, all I have to do is estimate the phase. Why would I not just coherently estimate the phase rather than do a differential estimation? I questioned this for a long time since there are a number of schemes to coherently estimate the phase.

SIMON: Yes, I will say by the way that there's another twist to the scheme and that you can actually use the first $N-1$ symbols to produce your carrier phase estimate and then do what we again think as classical differential detection which is related sort of to what you are suggesting.

HUTH: You see the advantage of coherently estimating the phase. I do not even have to use differential coding and therefore, I can eliminate its associated loss as well. I do not understand why I would want to use a differential estimation at all if I am going to take n symbols across the hop?

LINDSEY: My opinion on the answer to that question is as follows. If I understand what you're talking about, and I think I do, you do still have the problem of phase ambiguity resolution to take care of even if you go with the coherent scheme from hop to hop which means open loop phase estimation. I think its simpler, I think the algorithm that you're talking about is simpler, so that's why perhaps you would do differential encoding, to avoid that. You can avoid that ambiguity, I understand in frame sync and many other ways, if that happens to be part of your waveform design you would exploit that. The other issue might be dependent on how you do the algorithm from hop to hop in terms of getting the phase quickly, it might turn out that there's a hang-up problem, and I suspect strongly that there is a hang-up problem, that the repetitivity with which you gather and gain the knowledge of the phase is totally dependent upon what random phase effort, relative phase effort you come up with at the beginning of the hop.

HUTH: Not all phase estimation algorithms have a hang-up problem.

LINDSEY: As I say, it is algorithm-dependent. It depends on how you do it so I'd have to ... there's many ways. Proakis has some techniques in his books that I've looked at that are free of this hang-up problem.

SCHOLTZ: I think the communication theorists are seizing a session, finally. [LAUGHTER] I think we've given Marv at least a little feedback, probably what he was looking for from this group. So at this point, I'd like to turn the session over to Paul Sass here, who I think is going to give us a forward look into Army plans for communications.

PAUL SASS: *Issues in Tactical Networks*

Thank you Bob. I'm going to try and move kind of quickly although I won't be hitting you with any equations, so it should be fairly enjoyable.

What I planned on talking about was an effort that we started about a year ago at Fort Monmouth to try to make sure the communications side of the Army in the local area can keep up with the C^2 (Command and Control) side of the Army over the next 20 years. In fact this Tactical Local Area Networks (TACLAN) program that I'm going to talk about is a long-range plan to help us evolve from where we are today circa 1995 to where we have to be by the year 2015. So we are looking ahead very far. My plans to start the talk with a discussion of where we are today are supported by Seymour Stein's comments yesterday, so I'll begin with a real brief run

down on what we have in this 1990-1995 time frame in the area of tactical communications.

As we said yesterday, the Army is currently buying 3 major communications systems. The three systems are: the area system called MSE (Mobile Subscriber Equipment), the Data Distribution System called EPLRS (Enhanced Position Location and Reporting System) with JTIDS providing a small part of that data distribution capability, and of course the Combat Net Radio which you know as SINCGARS. In all its wisdom, the Army about 5 years ago concluded that EPLRS would provide data distribution for the Army. The other systems are basically voice systems although each of them has some limited data capability. I'll very quickly describe each of them for those of who haven't seen them.

EPLRS is the first spread spectrum network to actually reach the field. It's a time-synchronized, direct sequence distributed network that operates in the 400 MHz band, 420 to 450 MHz actually. Its primary purpose was to do position location among the network participants and to distribute that information as needed. It does it by doing time-of-arrival measurements, and passing those time-of-arrival measurements to a central computer called the Net Control Station which is not shown on this chart. That net control station computes, displays and distributes positions of the several hundred network participants. EPLRS was developed based on an Army marine requirement a long time ago. Five or ten years ago the Army decided it wanted to add some limited data capability to these spread spectrum communicating radios. So very limited data capability was added and about 1 kilobit per second of user data throughput can be supported between terminals. EPLRS supports data needlines that exist today. It is an AJ system, using direct sequence 5 MHz MSK signals, and hops over a small number of channels to increase the number of networks that can be supported in the same geographical area. It is fully secure, the security part being the biggest barrier to the development and the main reason it's taken so long to get it to the field.

Again as I said, EPLRS uses a synchronized, time-slotted architecture similar to JTIDS (many of you are familiar with JTIDS). It allocates time slots to specific users depending on their position update needs, depending on their movement on the battlefield. A manpack will get fewer time slot allocations than a moving vehicle. EPLRS consists of small user units (manpack radios) on the order of 500 per division, all controlled by one central net control station, or multiple redundant net control stations, but it is a centralized architecture. This is the picture of what the radio looks like. Again, as I said, it provides to the user about a kilobit per second of usable data rate. EPLRS is being built by Hughes Aircraft and I encourage you to discuss with them the applications of your theory to their system needs.

MSE, the Mobile Subscriber Equipment, is another system that the Army got very serious about deploying quickly. A few years ago the Army decided it couldn't tolerate the 15-20 year development cycle any more. The Army went out and decided to acquire a nondevelopmental (NDI) Mobile Subscriber System. In fact, the fielding dates have proven that you can buy a system off the shelf in fairly short time frame, much shorter than we've demonstrated in the past.

MSE is basically a circuit switched telephone system for the battlefield. It provides mobile subscriber services; it includes

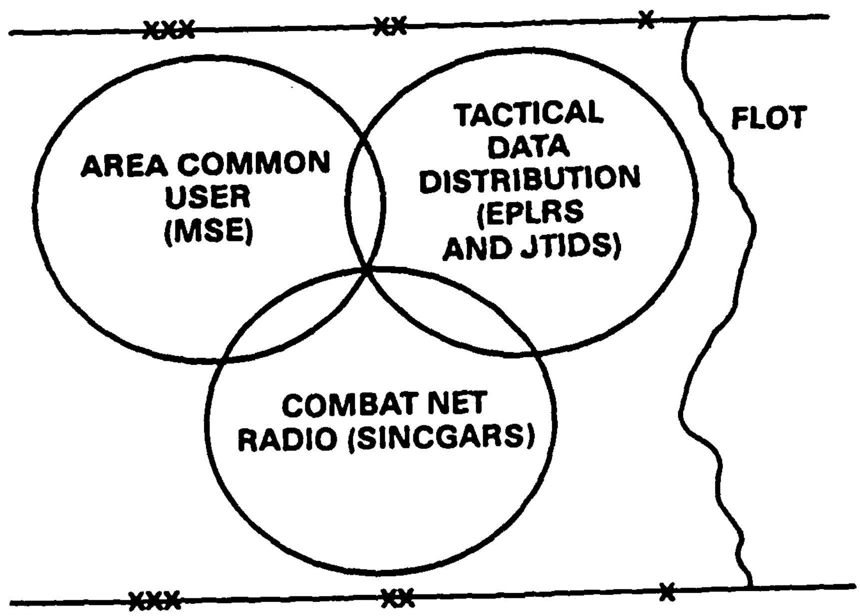
TACTICAL LOCAL AREA NETWORKS (TACLAN)

PAUL SASS
CENTER FOR C3 SYSTEMS
US ARMY CECOM
FORT MONMOUTH

CECOM CENTER FOR COMMAND, CONTROL AND COMMUNICATIONS SYSTEMS

EPLRS IS ONE OF THE ARMY'S TRIAD OF BATTLEFIELD COMMUNICATIONS SYSTEMS

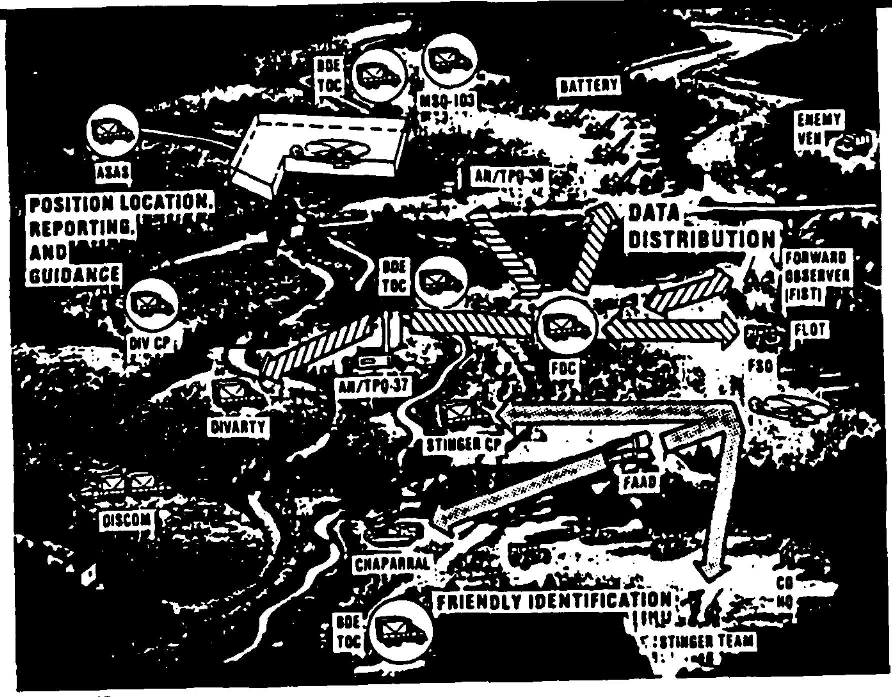
HUGHES



SLIDE #2

EPLRS MAXIMIZES BATTLEFIELD EFFECTIVENESS OF CURRENT AND NEW GENERATION SYSTEMS

HUGHES



SLIDE #3

REQUIRED OPERATIONAL CAPABILITIES

HUGHES

ENSURE EFFECTIVENESS OF EXISTING AND EMERGING BATTLEFIELD AUTOMATED SYSTEMS

DATA

- SATISFIES SPECIFIC SPEED OF SERVICE AND THROUGHPUT REQUIREMENTS FOR PRIORITY DATA USERS
- DEPENDABLE DATA COMMUNICATIONS
  - JAM RESISTANT
  - SECURE
  - CONTENTION REDUCED BY MESSAGE QUEUEING

POSITION LOCATION REPORTING

- COMPUTING REPORTING POSITION LOCATION
- POSITION LOCATION
- NAVIGATION
- FRIENDLY IDENTIFICATION

SLIDE #4

a mobile radio telephone that provides dial-up access for a VHF radio-equipped user on the battlefield, but the basic part of MSE is a battlefield grid connected by line-of-sight multichannel shots. All these radio links shown here are multi-channel line-of-sight shots, connecting approximately 40 nodes in a division. The different nodes in this chart are Node Central Switch (NCS), SEN is a Small Extension Node, and there's another one called the Large Extension Node. There are several different hierarchical switch sizes, all of them connected by multi-channel line-of-sight shots. There's also an extension to mobile subscribers through a VHF radio that we'll talk about real quickly.

MSE provides 16 kilobits CVSD, full duplex voice to all of its subscribers, the primary subscribers being digital telephones connected by wire to the switches. As I said, there's also the capability to connect 16 kilobit digital telephones through a VHF radio to the switches. The system was bought as a complete package including shelters, fuel tanks, all kinds of support, to actually support a fielded system right out of the box. It was a 4 billion dollar procurement over the last few years and the primary contractor is GTE.

These are the components of the MSE system as far as communications goes, and there are several different users. The users include voice users through a variety of types of voice terminals, facsimile, as well as some very limited data capability through the phone terminal itself. As far as radio goes, there are several different types. There's a line-of-sight multi-channel radio that operates in the UHF band, basically 220 MHz to 2 GHz in several different bands just like today's multi-channel radio. There is a VHF radio called the Mobile Subscriber Radio Terminal (MSRT) that is a non-hopping full duplex FM radio. And there are a couple of other special purpose radios like an SHF down-the-hill shot for remoting certain switches in certain applications. I've also shown a radio access unit which is actually an assemblage of eight of these MSRTs. So you have a stack of 8 VHF non-hopping radios providing a radio access unit which is the access point into the circuit switched system for these VHF mobile terminals.

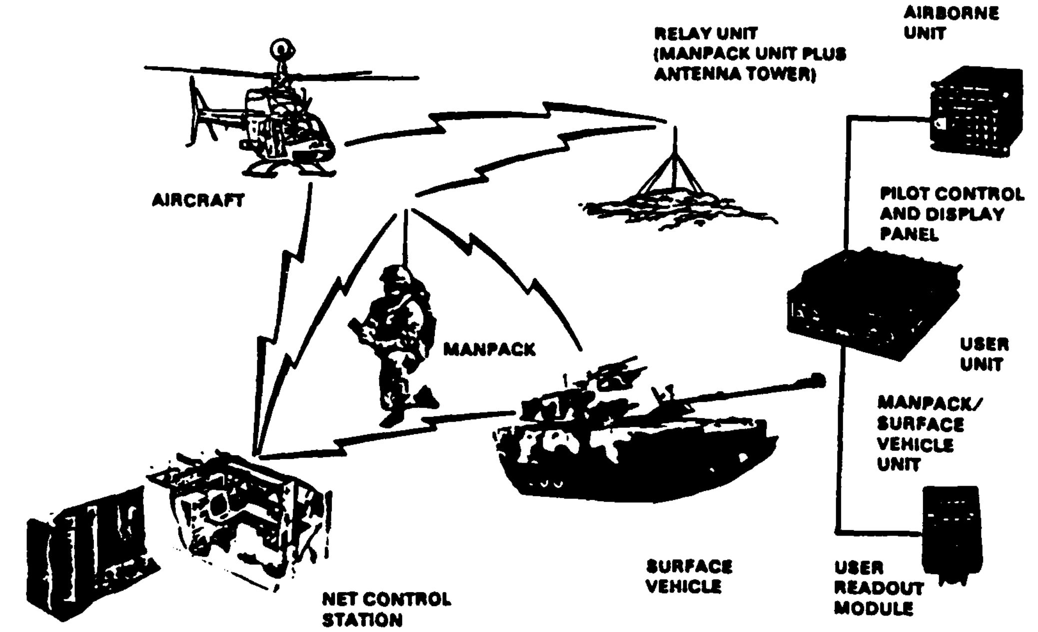
A couple of quick pictures. This is what a multi-channel radio antenna assemblage looks like with the large tactical radio. I am going to accelerate through these. As I said, there are several band heads in this UHF radio that cover about 220 to 2 GHz. The group data rates for these backbone trunk channels vary from 250 KBps up to 1 or even 2 MBps capacity for interconnecting the switches. This photo shows the MSRT, the mobile radio telephone for the battlefield that fits in a mobile vehicle providing the mobile telephone with access into the switched system.

I wanted to show you what a typical switching center looks like on the battle field. Now shown here something called a large extension node. Physically this is what makes up one of the nodes in this distributed network. It is physically a number of trucks interconnected with a lot of wiring, providing service to different numbers of telephone subscribers for each of the different size switches. Different size switches support different numbers of subscribers obviously.

One recent addition to the MSE systems is called a packet overlay. The Army just recently agreed to add packet switched data capability to this circuit switched system. We're going to overlay packet switches onto the backbone line-of-sight multichannel network just the way ARPANET or MILNET

EPLRS MAJOR ELEMENTS CONSIST OF NET CONTROL STATIONS AND ENHANCED PLRS USER UNITS

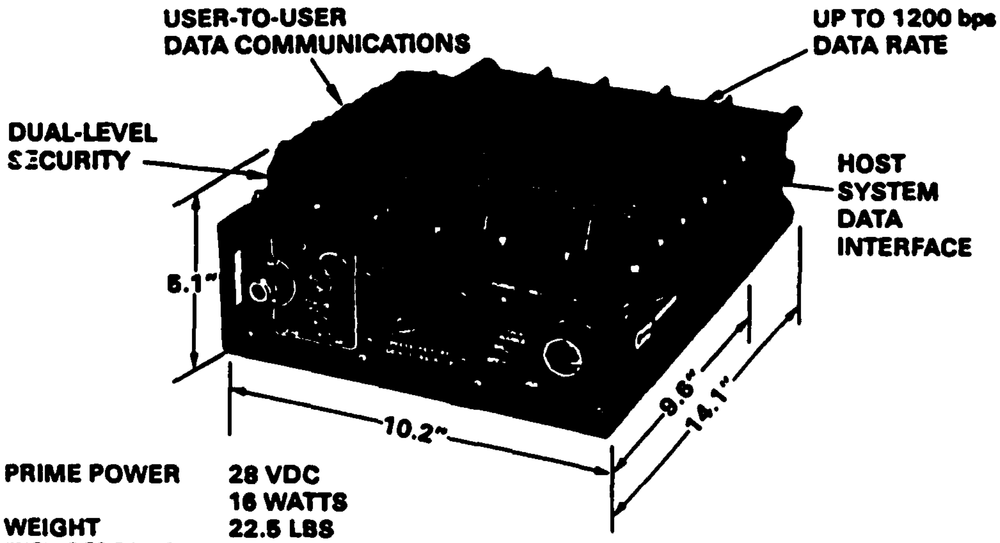
HUGHES



SLIDE #5

ENHANCED PLRS USER UNIT

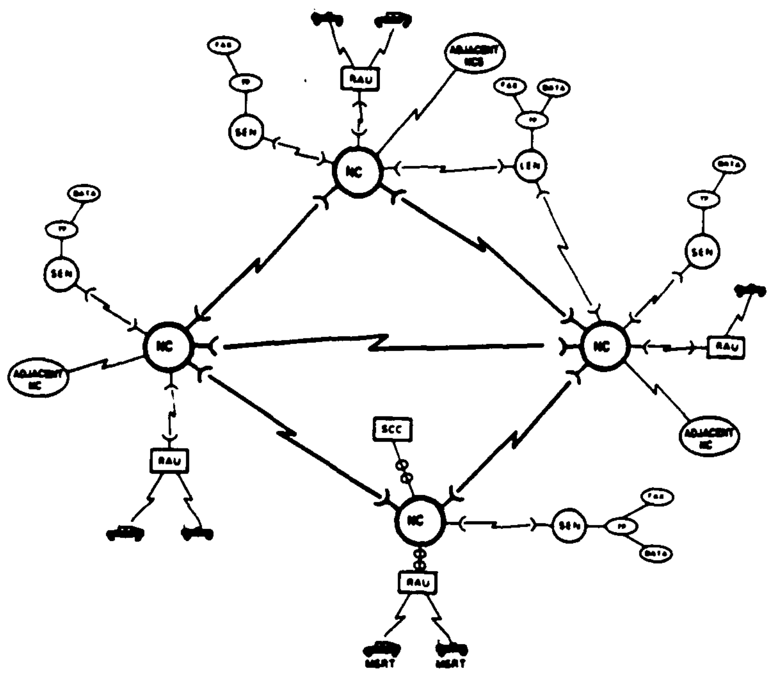
HUGHES



PRIME POWER	28 VDC 16 WATTS
WEIGHT (INCLUDES BATTERIES AND BATTERY BOX)	22.5 LBS
VOLUME	660 CU IN.

SLIDE #6

MSE Connectivity



TSD88-16

SLIDE #7

GTE MSE

MSE System Features

- 16 kb/s digital 42 node network
- Flood search routing
- Automatic transfer of subscriber affiliation
- Fixed directory numbers
- Deducible NATO numbering plan
- U.S. COMSEC
- Voice and data capability
- Downsized for increased battlefield mobility
- Provides roll on-roll off loading on C-130/C-141 aircraft

(CON'T)

TSD88-38

SLIDE #8

GTE MSE

MSE System Components

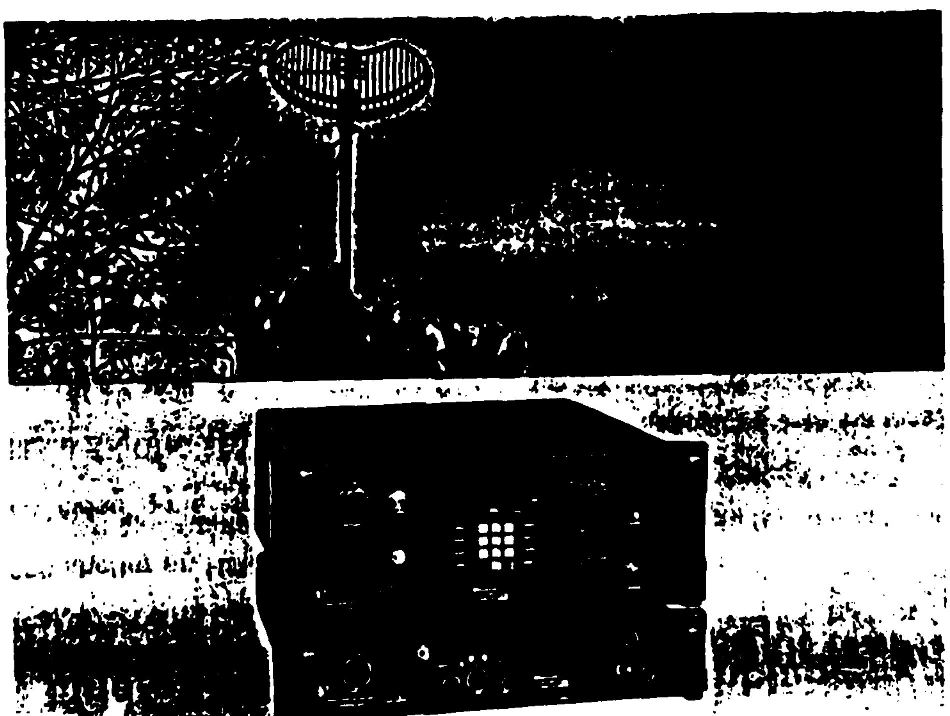
- **Users**
  - Mobile Subscriber Radio Terminal (MSRT)
  - Digital Non-secure Voice Terminal (DNVT)
  - Digital facsimile
- ***Communicators***
  - Switching
    - — Node Center Switch (NCS)
    - — Large Extension Node (LEN) switch
    - — Small Extension Node (SEN) switch
  - Transmission
    - — Line-of-sight (UHF) radio assemblages
    - — Down-the-hill (SHF) radio equipment
  - Mobile subscriber access
    - — Radio access unit (RAU)
  - Network control
    - — System Control Center (SCC)
    - — Management facility
    - — SSC interface

(CON'T)

SLIDE #9

GTE MSE

Line-of-Sight Radio



SLIDE #10

GTE MSE

Line-of-Sight Radio Set
AN/GRC-226

- **Frequency synthesized, microprocessor-controlled, digital radio**
- **Two frequency bands**
  - 220 to 400 MHz (Band I)
  - 1350 to 1850 MHz (Band III)
- **Three group data rates**
  - 256 Kbs
  - 512 Kbs
  - 1024 Kbs
- **Components**
  - Base band unit
  - RF Diplexer Unit

SLIDE #11

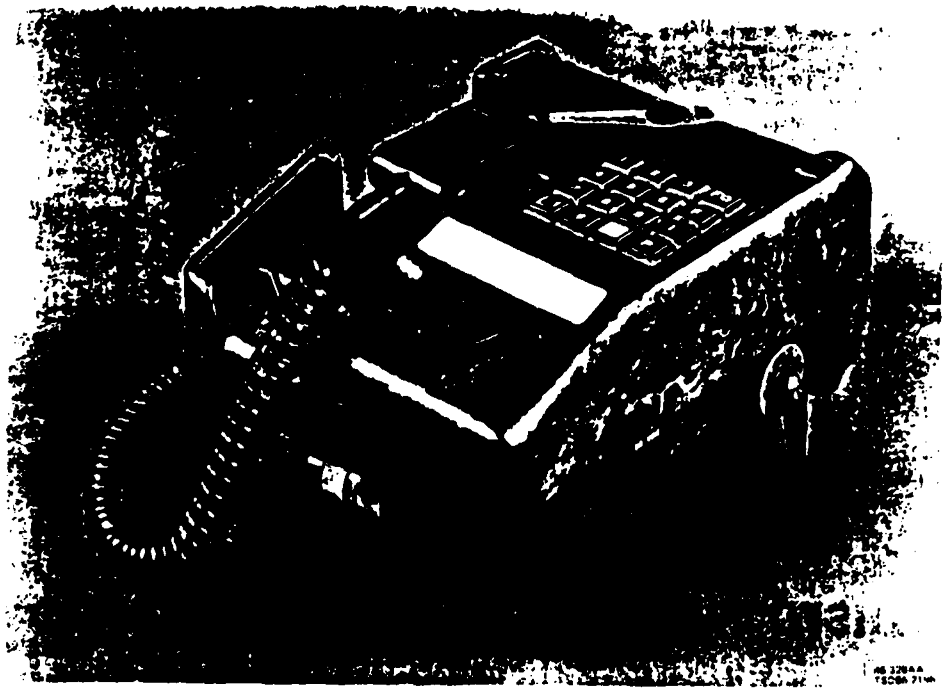
GTE MSE

Mobile Subscriber Radio Telephone Terminal

GTE MSE

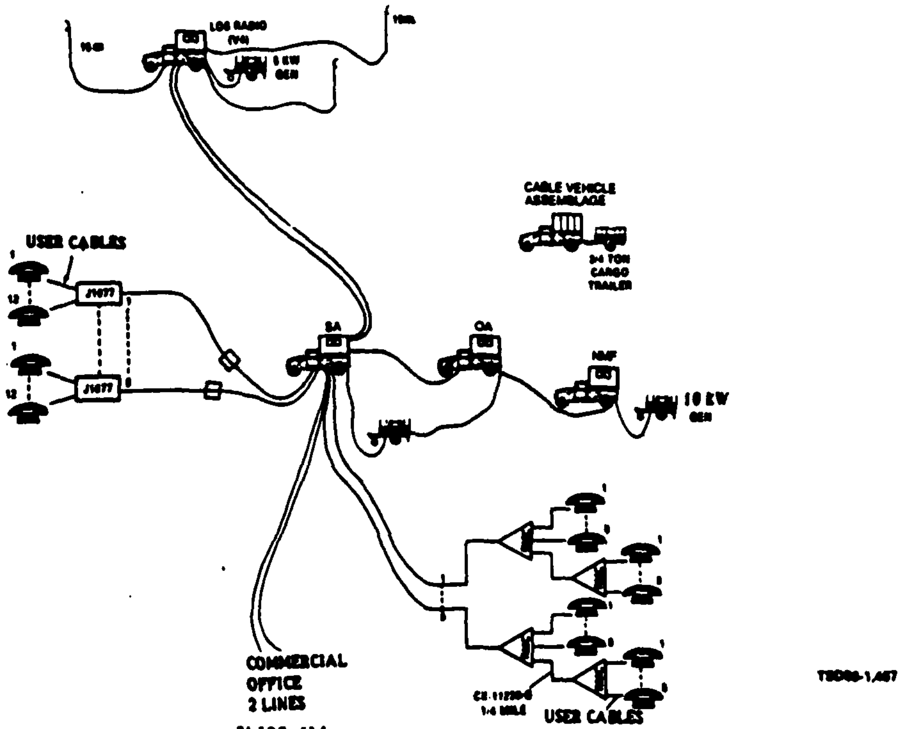
SLIDE #12

Digital Non-Secure Voice Terminal (DNVT)



SLIDE #13

Typical LEN Site



SLIDE #14

or DDN does its packet switching. In fact the packet overlay that was just recently purchased adds data service to MSE, using a standard BBN commercial product, a C3 packet switch, to access the trunk side of the switch. It connects at the trunk side of the circuit switch and provides ethernet as well as X.25 service to the subscriber. This is a very important addition that we've fought for about 7 years now.

The last of the existing communications that I'm going to talk about is SINCGARS. It's a slow frequency hopping 16 KBps voice radio. It's primary purpose is to provide voice communications on the battlefield among all the lower echeleon users. It also has a data port which supports limited data capability. It allows 16 KBps without any error control at all and with majority vote repetition coding it will provide integral sub multiples of 16 KB but it has very limited data usage anticipated, at least as of today. One of the things you heard Mike Pursley mention is that we've been working for a number of years on developing packet data capability for the SINCGARS radio. It's a recognition at least in part of the community that those that have the SINCGARS radio are going to want to build packet data networks out of the radios. To do that you have to basically add a network layer protocol above the link and physical layer that the radio provides. In doing that you also want to upgrade the link layer error control that the radio is providing. So that's in fact what we have been doing in the packet applique program. The SINCGARS radio is being built in quantities of tens of thousands and the two contractors today are ITT and GD.

O.K. that's all I want to talk about. That describes, in summary form, what tactical communications are like today. We will have it bought and we'll have the bills paid by 1995 as Seymour mentioned yesterday, and therefore we're going to start looking to what we are going to do now. I'll now describe what we see over the next twenty years.

As far as TACLAN, we started a year ago trying to conceptualize an image of where we wanted to be in the 2010-2015 time frame. We identified a real void in the battlefield. The void becomes very obvious when you start recognizing that data usage is going to explode. The Army is now buying a family of computers called the common hardware-software family which is a family of HP MIL-TOPE Unix machines that has a capability of being local area netted on the battlefield. In doing that the Army puts its Command and Control applications a step ahead of its communications. So we realized that we're going to have to get serious about providing certain things in the local access area, one of which is local area nets, and we've gone a lot further than that. We recognized that we can't afford a wired command post; we have to have the ability for survivability of moving command posts very quickly and not being constrained by the kind of wiring I showed you in that large extension node wiring diagram. So we need the ability to communicate with large capacities, wireless local area nets if you will, while on the move or just in between moves. We've got to get away from wire. Fiber, while it gives us the capacity, doesn't really get us away from the wiring problem although it does simplify the number of wires you need. That's the short term view.

The next chart discusses the midterm. Here I used an acronym, SICP (Standard Integrated Command Post), which is a program the Army has embarked on to basically construct a family of command post configura-

SINCGARS

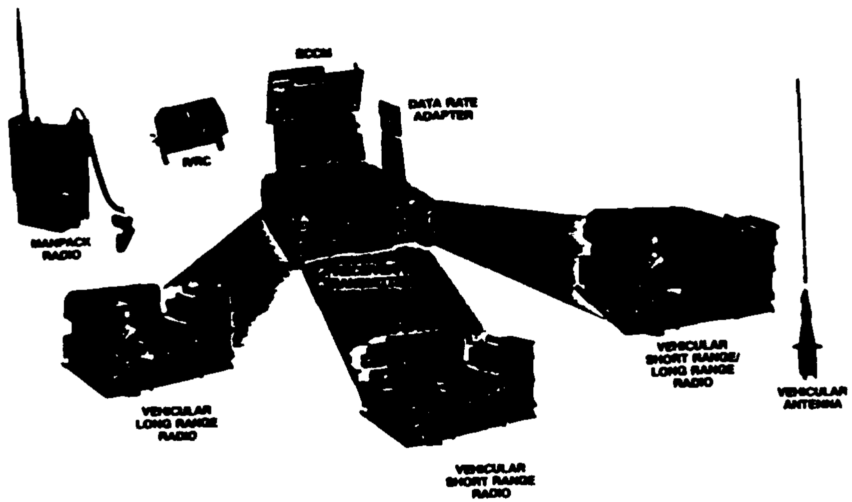
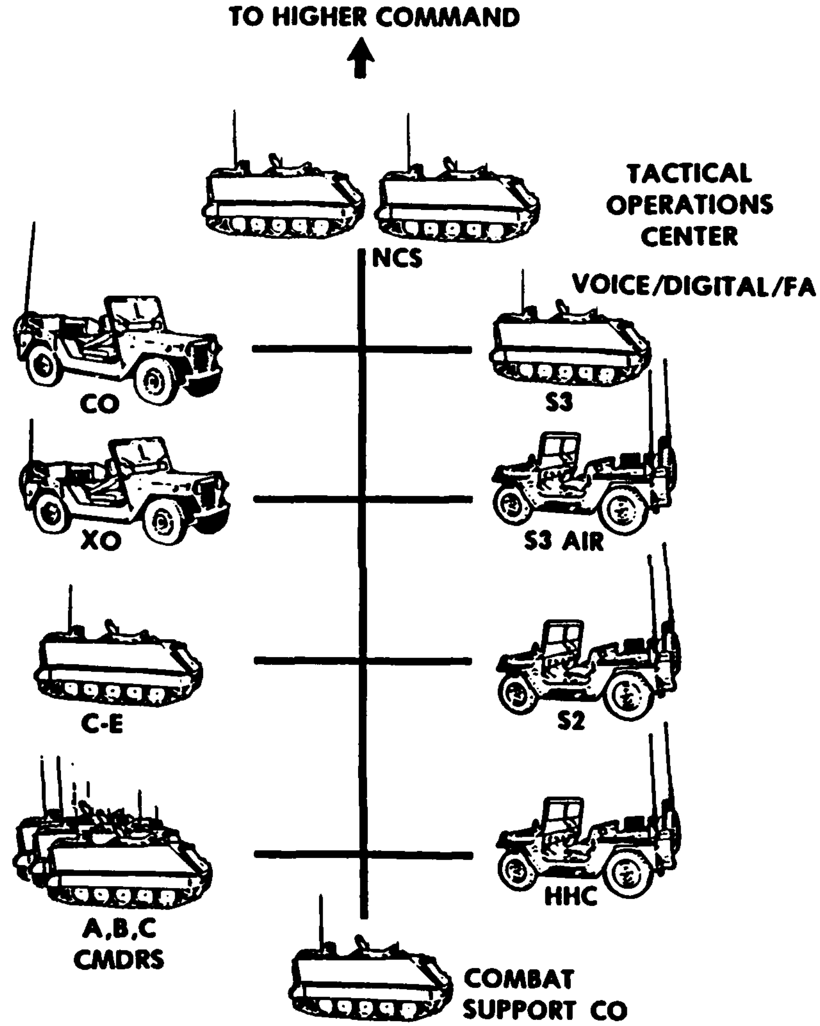


Figure 1. SINCGARS Collage

SLIDE #15



TYPICAL BATTALION LEVEL COMMAND/OPS NET

SLIDE #16

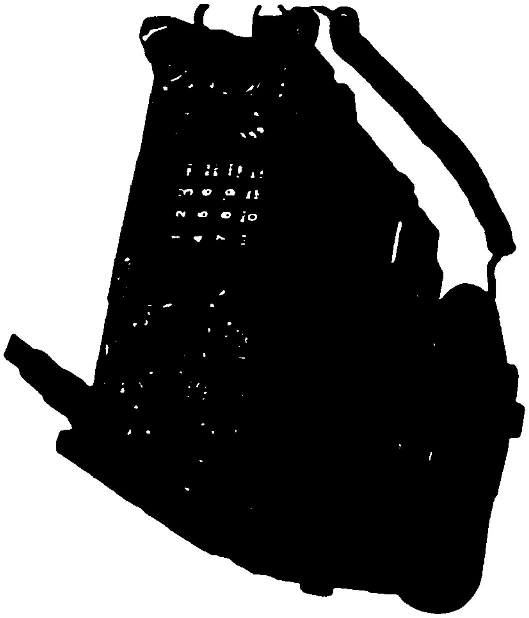


Figure 8. ITT SINCGARS Manpack Radio Subsystem

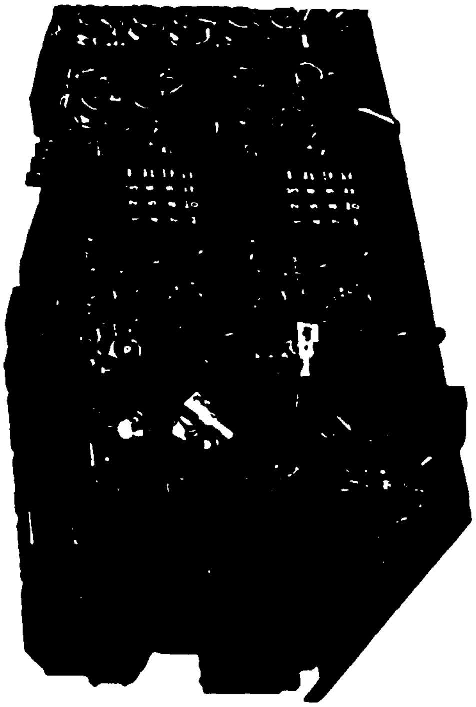


Figure 9. ITT SINCGARS Short-/Long-Range Radio System

SLIDE #17

WHY IS IT NEEDED?

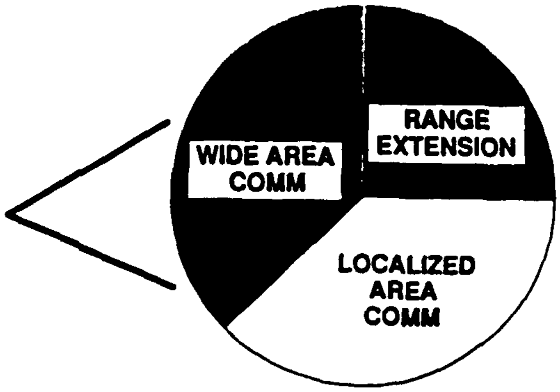
- **SHORT TERM VIEW:**
  - –DATA USAGE WILL GROW AS MORE BATTLEFIELD AUTOMATED SYSTEMS ARE FIELDED
  - –LAN USAGE WILL GROW AS WIDEBAND LANS LEAD TO MULTIFUNCTION (IE, VOICE/DATA) WORKSTATIONS
  - –LANS APPEARING WITHIN COMMAND POSTS (IE, SICP) AND AT VARIOUS ECHELONS WILL NEED HIGH CAPACITY CIRCUTS WHICH PRESENTLY DO NOT EXIST
  - –FOR SURVIVABILITY, CPs NEED WIRELESS COMMUNICATIONS, DISPERSION BY FIBER, AND THE ABILITY TO COMMUNICATE WHILE ON THE MOVE
- **LONG TERM VIEW:**
  - –MORE VOICE USERS WILL MIGRATE TO LOCAL NETWORKS
  - –ISDN WILL INEVITABLY IMPACT SUBSCRIBER DEVICES, NETWORK SERVICES, AND REQUIRED CONNECTIVITY
  - –IMPACT ON WIDE AREA NETWORK?

USER VIEWS ARE CRUCIAL!

CECOM CENTER FOR COMMAND, CONTROL AND COMMUNICATIONS SYSTEMS

SLIDE #18

TBIS INFORMATION TRANSMPORT



LOCAL AREA — **WIDE AREA**

INTEGRATED VOICE/DATA/IMAGERY
SURVIVABLE INFORMATION

MOBILE COMMUNICATIONS
LANs
HIGH BANDWIDTH
FIBER
ENHANCED CNR
MMW, IR WIRELESS LANs

GLOBAL INFORMATION FLOW
HIGH CAPACITY BACKBONE
PACKET AND VIRTUAL CKTS
DYNAMIC RECONFIGURABLE NETWORK

AND THEIR INTERCONNECTION
THROUGH
NETWORK–OF–NETWORKS

CECOM CENTER FOR COMMAND, CONTROL AND COMMUNICATIONS SYSTEMS

SLIDE #19

STRATEGY

- ENHANCE PERFORMANCE OF EXISTING/NEW SYSTEMS
- EMPHASIZE LABORATORY TESTS, USER TEST BEDS, AND TECH DEMOS
- DEVELOP MILITARY–UNIQUE ASPECTS
- TAKE ADVANTAGE OF:

–INDEPENDENT RESEARCH AND DEVELOPMENT PROGRAM

–JOINT PROGRAMS

–DEFENSE ADVANCED RESEARCH PROJECTS AGENCY

–COMMERCIAL, NON–DEVELOPMENTAL ITEMS

–CANADIAN AND FOREIGN DEFENSE SHARING IN R&D

CECOM CENTER FOR COMMAND, CONTROL AND COMMUNICATIONS SYSTEMS

SLIDE #20

tions out of a standard module, a standard shelter which has a bunch of computers local area netted together. We're trying now to address how that local area net will tie into the backbone. Over the long term we see ISDN becoming a reality. We are trying to capitalize on what the commercial world is doing in ISDN. I'm not sure how it will evolve into the access system. It may produce terminals down at the subsciber level, but how it will evolve into the switching and access system of the Army which is given to us by MSE at this point is not clear.

About a year ago the Army made a very concerted effort to develop something called a technology-base investment strategy. It was a long-term investment plan to attack the different communication problems and try to develop strategies for solving them over the long term. It basically divided communications into two regimes; local area and wide area. What I've tried to do here is summarize the key capabilities that we see being needed over the next 20 years in each of those areas. The key message is capacity; you see a lot of packet data networks being used to glue together all these heterogeneous communications systems in this notion of interconnected networks; the internetwork idea is being viewed as the answer to having these communications systems effectively interoperate on the battlefield. In the local area we see more use of higher bandwidth media whether it is fiber LANs or just ethernets on coax or wireless LANs. Low capacity wireless LAN capability, like an enhanced combat net radio, perhaps can provide, is included in the long term investment strategy. We also see that to achieve increases in capacity, and to get a lot of other things, we're going to have to move up higher in the frequency band. Personally I've pushed real hard for millimeter wave as the band we should target. There are a lot of nice attributes in that band and I think it's a perfect combination of packet networking with a propagation medium that is difficult. Those two technologies together can get us in the direction we need to go.

The strategy we've identified is largely limited by reality, reality in the way of budgets, and reality in the way our procurement system has worked over the past. We are going to start with an existing base of these systems as they exist today. We are going to be limited in budgets so we are going to have to leverage as much as we can all the things that are going on in the commercial world, in IR&D programs at private industry, as well as through joint programs among services. And we feel this internetwork architecture is a key area that's going to need joint development over the next 10 years among all the services.

This slide is a little redundant. As far as TACLAN goes, TACLAN is being conceptualized as the next generation local access system for the Army. It's got to start with what we have today, it's got to provide wide-band interconnection among mobile fighting units, and dispersed command posts. That's its major target. As LANs start to become more used, it will be a natural evolution to migrate voice and other services onto the LAN. That will help the wiring problem too. We see a number of technologies play a major role in this migration process. LAN technology coming out of the commercial world is already finding its way into military applications; both fiber and coaxial packet technology is going to be big part of the answer. That packet overlay onto MSE was a very major step in our view indicating an acceptance of packet technology as the glue that's going to tie together all these networks. There's a lot of gateway and bridge technology coming

WHAT IS TACLAN?

TACLAN IS THE ARMY'S NOTIONAL SYSTEM FOR LOCAL AREA COMMUNICATIONS

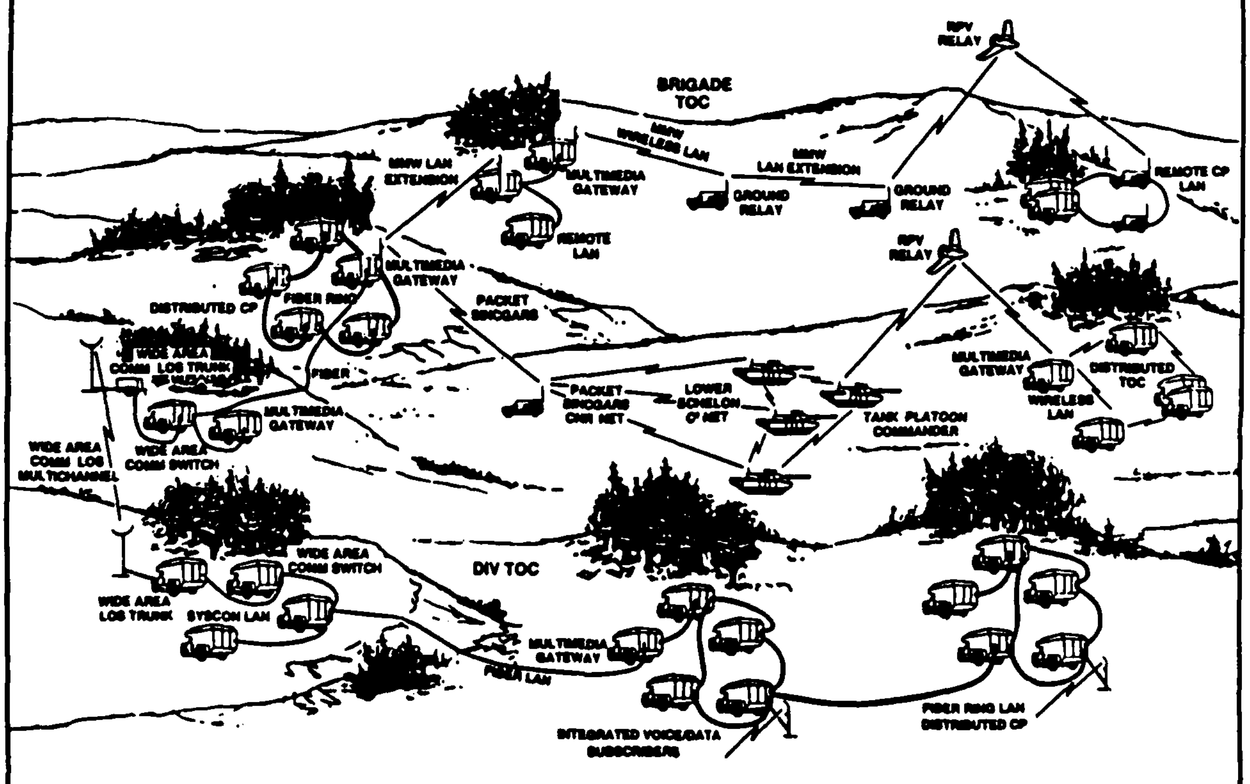
TACLAN:

- MUST EVOLVE FROM TODAY'S COMM ARCHITECTURE
- WILL PROVIDE SECURE, SURVIVABLE INTERNETWORKED COMMUNICATIONS AMONG AUTONOMOUS FIGHTING UNITS AND DISPERSED CPs
- WILL INTEGRATE DATA, VOICE, GRAPHICS, AND PERHAPS VIDEO
- IS AN APPLICATION WHICH INTEGRATES:
  - –COAXIAL AND FIBER LAN TECHNOLOGY
  - –PACKET TECHNOLOGY
  - –GATEWAY AND BRIDGE TECHNOLOGY
  - –WIDEBAND MMW/MICROWAVE RADIO

CECOM CENTER FOR COMMAND, CONTROL AND COMMUNICATIONS SYSTE

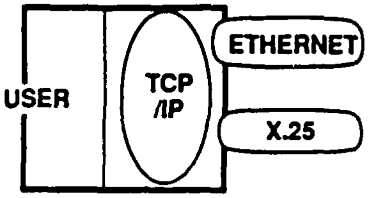
SLIDE #21

TACTICAL LOCAL AREA NETWORKS (TACLAN) –2015–

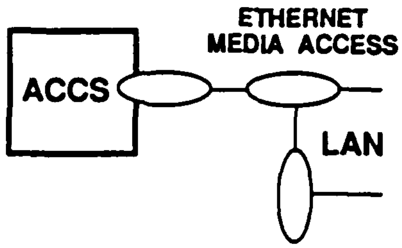


SLIDE #22

WHO ARE TACLAN'S SUBSCRIBERS?

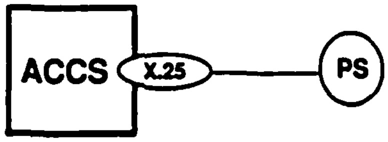


ACCS HP9000/300 SERIES



ACCS WITH LAN CONNECTION

ACCS STANDARD HOST SUPPORTS X.25 AND ETHERNET LAN CONNECTION AND A FULL SUITE TCP/~~TCP~~ IP IMPLEMENTATION



ACCS WITH X.25 CONNECTION

SLIDE #23

CECOM CENTER FOR COMMAND, CONTROL AND COMMUNICATIONS SYSTEMS

HOW WILL TACLAN HELP LOCAL AREA COMMUNICATIONS?

- WIDER DISPERSION OF CP SHELTERS
- MOBILE LAN OPERATION
- LESS WIRING/FASTER SET–UP & TEAR–DOWN
- VOICE/DATA INTEGRATION

FLEXIBLE LAN INTERCONNECTION

SLIDE #24

CECOM CENTER FOR COMMAND, CONTROL AND COMMUNICATIONS SYSTEMS

TACLAN

- **MIDTERM GOAL (NEXT TEN YEARS):**
  - WIDEBAND WIRELESS/FIBER INTERCONNECTION OF LOCAL AREA SUBSCRIBERS INTO ORGANIC TACTICAL COMMUNICATIONS
- **COMPONENT CAPABILITIES:**
  - SECURE VOICE/DATA LAN
  - WIRELESS LAN TECHNOLOGY
  - FIBER OPTIC LAN/LAN INTERCONNECTS
  - TACTICAL MULTIMEDIA GATEWAYS
- **APPROACH:**
  - PHASED DEMONSTRATIONS
  - INTEGRATION INTO ATCCS

SLIDE #25

CECOM CENTER FOR COMMAND, CONTROL AND COMMUNICATIONS SYSTEMS

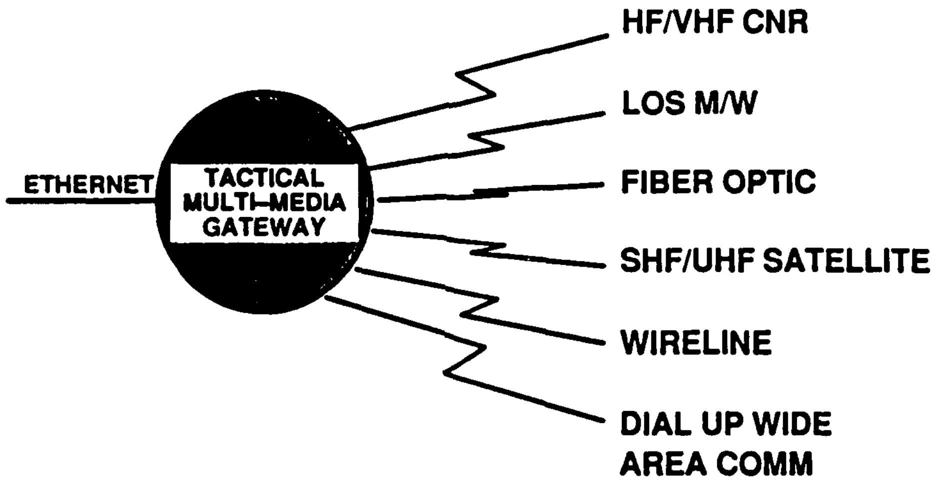
TACLAN COMPONENT TECHNOLOGIES

- **INTEGRATED VOICE/DATA ON SECURE LANs**
- **WIRELESS LANs**
- **FIBER OPTIC LANs/INTERCONNECTS**
- **TACTICAL GATEWAY TECHNOLOGY**

SLIDE #26

CECOM CENTER FOR COMMAND, CONTROL AND COMMUNICATIONS SYSTEMS

OBJECTIVE
MULTI–MEDIA INTERCONNECTION



SLIDE #27

CECOM CENTER FOR COMMAND, CONTROL AND COMMUNICATIONS SYSTEMS

out of the commercial world to interconnect local area nets and, as we've said, the move higher in RF frequency to get higher capacity interconnections is something we are anxious to push. I sat down with the artist to try to come up with a picture and this has a lot of things in it. It's mainly useful for talking to and I don't have a lot of time to talk to it. The technologies that are shown here include a variety of media to interconnect local area nets. Now, the local area nets are not only within single shelters. The local area nets start to leave controlled areas connecting together different groups of shelters. In fact, in most applications you see a command post consisting of a bunch of vehicles dispersed over as wide an area as they can. The wider they can disperse, the more survivable they are in a tactical environment. That creates a lot of problems. You see fiber interconnecting a lot of them, you see a number of applications. I've shown a wireless local area net where you have an application where you don't particularly want to wire up or perhaps you want to interconnect mobile vehicles, as in tank-to-tank applications. There's a lot of use of relay extensions as Seymour mentioned. We're talking about packet networks very largely to interconnect distributed heterogeneous media. We have RPV relays that can interconnect an isolated local area net sitting out here in the boondocks with a command post through various different media. Motivation for fiber is capacity, not necessarily interception as much, that definitely helps, but ....

DENNIS HALL: What kind of rates are you looking at to justify fiber?

SASS: The fiber systems we are using now are looking at FDDI. So we are going to 100 MB or even higher capacities on the fiber nets. We are also looking for survivable fiber architectures. So a single ring fiber topology isn't really going to hack it in the long term. We need survivable dual-ring architecture like FDDI-2. That's the long term. Initially just a fiber ethernet will satisfy a lot of the preliminary capability in the 1995 time frame. The local area nets the Army is buying will have a fiber capability to interconnect into MSE.

HALL: 10 MBs on the ethernet?

SASS: Yes 10 MBs on the ethernet as the start but as you start to move more of the voice users on to the data LAN, then you start needing more bandwidth. We don't yet have a good feel for the load, the capacity required, but the evolution that we see happening is as more LANs are installed for data users, the next logical step is to move those voice subscribers onto the land. Once you start doing that you start chewing up capacity. That's when you start needing more and more bandwidth.

CIOFFI: Following onto his question at the 100 MB rate on the fiber, what lengths are you looking at for distribution? How many kilometers?

SASS: I can't answer that right now. The initial deployment of fiber LANs is a capability of distributing a command post I think over an 8 kilometer circle with a number of nodes in that 8 kilometer. It's that order of magnitude, it's kilometers that we are looking for. That isn't supported by the 100 MB FDDI right now.

Let me talk real quickly about what some of the pieces are. What this TACLAN notion is going to do for us is summarized on this chart. It will let us distribute these command posts over larger areas than they can right now. Right now they are wired together, initially they'll be wired together either on coaxial or this initial fiber implementation that I told you, an 8 kilometer ring. That's not even

wide enough. For survivability in the battlefield you'd like to distribute your command post over even wider ranges. What's more important though is you'd like to have the capability to move your command post frequently as well as while operating. So therefore a wireless LAN capability becomes more important. Right now our only wireless communications is very low capacity, and it's very questionable whether we can use something like a packet radio which is an experimental device, a couple of 100 kilobits to adequately interconnect mobile LANs. That's one of the things we are trying to exploit over the next few years to just assess how fairly constrained throughputs can interconnect tactical lands doing an actual LANs in support of an actual C^2. And as the command and control functions develop we'll start to see what kind of throughput we need to interconnect them. Obviously if you move away from wire the better you are in terms of setting up and tearing down a command post. The argument about voice and data integration, I think it's natural and something that is going to take a lot of serious thought, once we put these data LANs in, I think the natural migration will be to get rid of all those single subscriber voice-telephone hookups you saw in the MSE wiring diagram. If you can integrate those telephones on to the LAN then you start significantly reducing the wiring burden of the command post which has computers and telephone subscribers sitting right next to each other. And what we are looking for is basically a flexible family of technologies that can provide interconnection in the tactical area.

We've targeted this a little better over the next 10 years. We're trying to focus a little more specifically on parts of that artistic picture I showed you. The specific areas we're going to be concentrating on over the next 10 years are wideband wireless and fiber LAN combinations, interconnection of the subscribers into one big organic local access system. The technologies that we hope to exploit are listed here. Number one, we need secure combined voice and data local area net capability, you need wireless LAN capability, we need a family of fiber LAN ... LAN-interconnection technologies which, if FDDI becomes real, will lead to FDDI, but I'm really not sure I believe that FDDI will ever be there. So it exists. I know it exists. I've heard some arguments in the community whether or not it will really be exploited in the commercial world. And that's what we are trying to leverage. We are trying to exploit whatever the commercial world picks up and carries.

And the last one is the notion of a tactical *multi-media gateway which I want to talk* about very quickly. The idea of a multi-media gateway I think is the last issue I'm going to talk about. Down at the subscriber, at this local area net, we would like to have a capability of giving the subscriber of command and control system the choice of using whatever media he has available. He may have an EPLRS network, he may have a SINCGARS network, he may have an HF network, and he has to have that choice of picking whatever media has to be in existence at that point, happens to be most efficient for his need. To do that we feel a gateway capability will be required down at the local area at the subscriber. Currently there is a device which the Army is buying which is a very limited capability in this direction and we hope to expand it to provide full ethernet gateway capability to a number of media, the media being links, the media being networks, the media being dial-up wire connections, whatever happens to be there to give the tactical user the most

flexibility. That's all I have.

SCHOLTZ: Are there any immediate questions? Yes John ....

JOHN TREICHLER: This is a simple one. I didn't see any potential application of VSATs (Very Small Aperture Terminals) or anything like that. Do you anticipate any use of things like that in the tactical setting?

SASS: I think I had a Tactical Satellite (TACSAT) as one of the media on that last chart. We in CECOM don't deal with the satellite community very extensively. SATCOMA is the Army's communications agency and of course their answer will be a lot different. It plays a part, it's got to be there as a media of choice for some applications. It does not fit in very heavily to what we are trying to show here in the local access area, but it certainly exists.

TREICHLER: The reason I asked is that it certainly seems to have pluses and minuses but a lot of the things you were very interested in moving toward, like no physical connections between the terminals and the ability to work on the run and things like that, seem to be things that VSAT was really built for. I mean if truck drivers can use them, then tank drivers ought to be able to use it too. I can think of a lot of disadvantages like survivability if you have only one or two relay points. If they are not survivable then your system isn't.

SCHOLTZ: I think Bob Peile has a question here.

ROBERT PEILE: I have two comments and a question. I worked on the Australian equivalent of MSE some years back. My first comment is that the Australians adopted fiber optic cable for 512 kbit/s. This replaced something like 7 tons of copper cable in an average HQ and replaced 2-ton vehicles with land rovers. It was no competition in terms of deployment flexibility. Capacity is another good reason. The second comment is that a lot of smaller nations, Norway for example, thought that the best upgrade to an MSE-type system was to put some switching capability into their multiplexors. If a switch disappeared, they couldn't talk to each other via a dumb multiplexor. An immediate upgrade was to distribute some very limited switching capability into the muscs. I think that's bound to continue. Finally, my question: The last time I talked to MSE people, they had an upgrade program within the existing MSE project for taking a few of the more glaring deficiencies out of the system they had bought. When you refer to 1995, is that something different than this upgrade program, or is that the upgrade program?

SASS: The upgrade program is part of the fielding schedule. What I was specifically talking about was when MSE would be deployed. In fact I'm not even sure 1995 is the correct published date for that.

PEILE: What you're talking about is outside of the upgrade program?

SASS: That's correct. The packet overlay is something that's being bought today and is going to be fielded with MSE by that 1995 time frame. There are other upgrades that the project manager plans throughout the procurement. I don't know whether it overlaps beyond 1995. I did want to comment about fiber. There's a lot of opportunities for fiber insertion into MSE and we are trying to push for that but as I said, MSE was bought as a complete package. There are a lot of rooms for improvement that we see. If fact there's a lot of willingness to now think about, now that the bill has been paid basically. And one very strong area is ECCM. There's a lot of opportunity both in the VHF mobile subscriber link, as well as the back-

bone line-of-sight.

LINDSEY: Paul, would you sort out the notion of what you mean by improved data rates. I heard numbers like 10 MB to 100 MBs, point one, and would you also elaborate on when you use the word survivable, what ingredients of that term is with regard to how you used. Finally, a further comment on, you were proposing millimeter wave bands, yet in the TACLAN drawing that you demonstrated all the tie-ends of various lengths of what not, I didn't see a satellite there but you did have a deployed mobile post that seemed to be a command post lying around. Was that the length that you were going to employ the millimeter wave technology?

SASS: You did see a satellite?

LINDSEY: I saw something up in the air, I don't know if it was a satellite or not?

SASS: It was an RPV relay.

LINDSEY: Oh, is that what it was!

SASS: There are a lot of opportunities for unmanned aerial vehicles as relays of opportunity. In particular, millimeter wave system might benefit from that. I'm trying to keep track of the questions.

LINDSEY: Point one was data rates.

SASS: What the Army has purchased with its family of new computers is a 10 MB ethernet capability as well as a 10 MB fiber capability along with that. The only reason for looking at things like FDDI is increased capacity as well as increased survivability from different ring architectures. I can't say that we've estimated the traffic load and we'll need 49.2 MB. We don't know yet. The only thing that's clear is as voice subscribers start to migrate on to the LAN, it chews up capacity rather quickly. I think each full duplex voice conversation is estimated at about 100 KB of capacity. So that starts to eat into your capacity and we'll need more.

LINDSEY: The next question was on survivability.

SASS: Survivability refers to the ability to keep communicating in the face of battlefield dynamics; dynamics caused by hostile action, EW threats, and mobility. The dynamic and severe nature of anticipated EW threats, as well as tactical mobility, are the two pressing requirements that present the greatest challenge.

One of the problems that we face is that we must advance technology by leveraging the developments in the commercial sector. But in most cases this doesn't necessarily give us the needed increase in survivability. We need to specifically target the issue of survivability, because it alone is the factor which differentiates our needs from those of the commercial world. So at the same time that we attempt to leverage commercial products, we need to apply our limited R&D resources to improve survivability and still get a capability into the field in a reasonable time frame. That's a challenge for all of us.

SCHOLTZ: Mike?

MICHAEL PURSLEY: Paul, you seem to minimize the data capabilities of SINCGARS, but it's considerably better than EPLRS. What is going to handle the data if SINCGARS isn't?

SASS: EPLRS is the Army's data distribution system ... [LAUGHTER] by decree. Nevertheless, there is an increasing awareness that SINCGARS will be the only radio that many of the data users will have. I don't agree that it is significantly better than EPLRS, but it certainly offers comparable capacity. At 16 KBps SINCGARS obviously appears better, but by the time you start to build a network and add decent error control, you're going to be down to a competitive kilobit or so.

PURSLEY: You are talking about network throughput rather than raw data rate out of the radio. But you're going to have the same problem with EPLRS. The raw data rate of EPLRS will also decrease when used in a store-and-forward network.

SASS: That's basically true. Let me clarify the question. EPLRS provides an advertised per circuit data capacity of up to 1200 baud simplex, or 600 baud full duplex. This is only a portion of the total capacity available to an individual EPLRS User Unit (EPUU). An individual EPUU can access a total capacity of nearly 4 KBps, much of which is consumed with overhead and, as you point out, depends on relay utilization.

Again I emphasize that there's been increasing awareness that the other systems are going to have to share the data load on the battlefield, as evidenced by the packet overlay onto MSE and the increasing interest in packet appliques to SINCGARS.

SCHOLTZ: If there are no other immediate questions I'd like to pass on to the next phase of this which is sort of a wrap-up session and some of the things we've talked about up to now including the last talk will probably pop up again somewhere in the wrap-up.

One of the things that I think would be very useful to our sponsors, the Army Research Office, is some discussion of where big payoffs might be in basic research in the various areas that we've been talking about. It will help them probably lobby for funds and I don't know what else they use these for. Bill Sander could probably even brief us on that. I've asked each of the panel chairmen perhaps to make a statement that we could chew over for 10 minutes or so each on where basic research opportunities lie in the four topical areas we have discussed. Do I have any initial volunteers? Bill Lindsey waved his finger. [LAUGHTER] Oh, is that where you're going Chuck or do you want to start it off? Bill, why don't you take the floor. After all, know your channels is first command and right, so let's start there and see what we need.

LINDSEY: It is the kind of a question that we are all interested in but I have to say that the answer is totally related to the kinds of things I'm trying to do in an analytical way.

I was involved, as you recall, with the channel question here and I guess if I look at this chart which tries to explain the various passages and paths that the topics we've looked at have addressed in the last two days. Relating that to the future, I guess I would tend to say that channels remain to be a problem of great interest in terms of their characterization. I would like to emphasize, however, the paths in this tree which I feel, if we investigate from the point of view of efficient communications, will render something in terms of benefits to the community. There are paths in this tree where I think perhaps we've reached the laws of diminishing return, i.e. to say, "Sure, there are interesting problems here to exploit and papers to publish." But what we gain back from it is not so much in terms of engineering knowledge. In particular, the path to the left here, what I call spatially invariant channels or perhaps frequency non-selective, and in particular time-invariant channels in the additive white Gaussian noise channel. This path has been beaten through to completion as far as I am concerned, whether it be NASA-related, commercial or military communications. In the telephone channel I'm not as familiar so I cannot say what is going on there and I'm sure there are experts here who can address that ... but if I had to put dollars into anything as far as channel

characterization understanding is concerned, this path I wouldn't proceed to exploit. This path again is time-varying, slow and fast; this is the case that I think we understand reasonably well from a modulation-coding point of view and there are some things that can be exploited there, but the big payoff in terms of investment of dollars is in my opinion not there.

Over on the other side, having to do with spatially varying type of channels where there are frequency selective effects, there are two paths listed there. The time-invariant channel is one we've talked about equalizing yesterday, but because of its relationship to the space-time channel we've talked about, that channel continues to need work because the techniques down this path happen to do with equalization. The solutions to problems along that path also fold in to techniques that can be used in channels which are slowly time varying as in the case of (HF) radio frequency type communications, or channels such as the laser channel (with clouds) is an area that continues to need work. So if I were to put red and green with regard to future investigations I would sort of shade this dark red, don't do anything there, I'd shade this little bit lighter red (I'm color blind so I couldn't tell the difference), then over in this path, certainly I would put green here. Green should continue in that direction in support of equalization, modulation coding techniques of the types that we've even heard here in the last two days. Those paths I feel are quite fruitful. Over in this direction, I think I'm of the opinion that there's a lot to be done, in particular, that's the path that I'm currently most interested in with regards to incorporating the physics of the medium into the channel model such that parameter sets associated with performance evaluations doesn't require one to say, well, if you have this parameter or preconceive what the parameter value should be. The effects of the earth's magnetic field, the effects of turbulence in the ionosphere can be folded into our channel models via of this invariant bedding technique that I talked about and I think that there's a lot of work that could be done there. In the slow case, this notion of adaptive equalization is going to be paying benefits as we learn to do that. That's the notion dealing with adaptive equalization in the spatial domain. Hence modulation and coding techniques are there to accommodate to channel models, or areas where I think are fruitful. And indeed in many applications there's knowledge that could be of value there. I think I'll cut off here and let the others amplify with respect to their own individual areas.

SCHOLTZ: I'd like to make one comment strictly from the academic point of view. That is, there seems to be a wealth of information out in this world somewhere. Maybe it's all in Phil Bello's head, or Al Schneider's or something like that. It would be a tremendous service to the academic community if some of these people would do a little digesting and put out a core dump in the form of a nice thick book which would really summarize their career experiences for us. At least I don't know of any book that communication theorists can jump in to and get really basic channel information that's deep and thorough. I would encourage anybody with that sort of background to please do something for us. That's my one comment to your summary, Bill.

LINDSEY: Thanks for reminding me. I believe that is a good comment. Also the notion of documentation is always an issue in this business, capturing knowledge and documenting it for those who go behind us and

with us and in front of us is always an issue in terms of research and cost. So frequently I see lots of knowledge being gathered but it's not being captured. How to fix that or ....

SCHOLTZ: ... or digest it ....

LINDSEY: ... digest it and capture for later use by those who go behind us or with us or in the front of us.

SCHOLTZ: Are there any more comments on Bill's summary in the channel area? ... I think we took the easy one first, let's see what comes next. Chuck, would you like to be next? Dennis is still working on his viewgraph over here. I know the guys you work with haven't put anything in writing, they don't speak up, I know the types ....

CHARLES WEBER: Yes, that's my first comment. After putting up my first couple of viewgraphs on Monday and listening to the speakers, I was left with the impression that many problems had been solved. Even though the efforts were algorithmically driven, few were written down. So, in order to get a feel for what people thought are academic issues that could be addressed in academia, I went to the speakers and some non-panel attendees. And, I think I ended up with more problems than were ever solved. I'm not sure I know where to begin. Some of the problems that I initially thought were only academic and not very pertinent, in view of all the algorithms that we have for modulation characterization, really are issues. There was a lot of consistency in the problems that came back.

The first was the area of bounds on modulation characterization performance. We have so many algorithms out there and so many things being done, there doesn't seem to be a benchmark as to what we might be striving for. How good is good? What really is good seems to be evading us. It wasn't clear from the feedback I got whether using maximum likelihood or generalized maximum likelihood or multiple classifiers, or something else is the approach that might be taken in order to create some goals that we can strive towards.

Another area is architecture. What architecture might be taken? One is what we were saying on the bounds, the maximum likelihood, generlized likelihood approach, as a methodology towards an architecture. The other suggestion was: should it really be a tree-oriented decision mechanism? That was left as a question whether each of these or both might be the approach to take. Maybe part of this is where neural nets, artificial intelligence, and training mechanisms fit.

The third generated most of the feedback, namely the area of improved algorithms, i.e. keeping on with what we are doing and extending it to sequential classifiers. I think Steve had several questions about sequential classifiers. Should they be used and what feature ordering is best? Is there a particular ordering of the chosen features and how robust are they? Is there a minimal set of features? If we have a criteria for performance, what is it? Is it minimum time, minimum observation period, best detectability, and does that go back to the bound because all those might affect the bound. It sounds like there's a gap between the philosophy of approach and the algorithms there. Maybe that's really what we are going after, maybe it's the missing link in all of this.

Measures versus features! Features, as I understand it, are like looking for something, then you see this feature and you say it's BPSK without knowing what the data rate or the carrier frequency is, but you know it's BPSK. You then go from there, as opposed to measuring these parameters up front. Al-

ternatively we may measure chip rate, not knowing whether or not it's BPSK. Is there something to be said between these two approaches?

Roger Peterson had a series of research questions, primarily involving wideband signals: determination of hop rate in strong interference, and at low SNR. It's true that a lot of the detectability algorithms for frequency hopping and/or direct sequence work well at intermediate signal-to-noise ratios, so when you get down to -5dB per chip or lower, then you have to integrate for a longer time. It seems that we need to know some waveform parameters a priori in order to be effective in determining the presence of wideband waveforms.

The question that Roger raised which was new to me, and an interesting one, was determination of the frame rate. Not only do you not know the code but you also do not know the frame rate, and you'd like to make those determinations. These are legitimate issues both in frequency-hopped and direct sequence environments. And also at in low signal ratios! Do we know how to do this at high SNR's?

SCHOLTZ: Since I don't really work in this field and Chuck is just starting, I guess I can make my two cents known here. Some of the problems here remind me of things that we used to do back in radar. I had a fair amount of radar experience back in the 50s, 60s and 70s. We typically started out designing a system and then we started thinking about countermeasures, and then we started thinking about counter-countermeasures and you keep playing this game and pretty soon you realize that the best counter-countermeasures are really just good system design procedures the first time around. I'm wondering whether somewhere in this community shouldn't be doing things more like, say, Mark is doing where you're studying what defeats the signal intelligence people and putting that in front in the systems you're building. And I haven't heard much discussion of any of the things that Mark was talking about, for example, in battlefield communications. Maybe that's too short term a lifetime to worry about. I don't know. But we are still talking about BPSK and QPSK. I haven't heard anything about some of the shaping things we're talking about and hiding lines and stuff like that. Maybe this is a question to Paul Sass to some extent. Are you really worried about intelligence in any of the networks that you are constructing and is that an issue that any basic research should be devoted to?

SASS: This is a tough question. It is clearly of concern and it plays a part in the developments as we progress.

SCHOLTZ: Should we be spending time on that for you in any way, or don't you even want us to talk about it?

SASS: No, I honestly don't know the answer. Let me just say that for a number of years we were approaching the design of an LPI distributed spread spectrum network. That was a pure thrust we were taking as part of what motivated a lot of the research efforts we funded. Because of the funding realities I've been describing, we weren't given the freedom of actually pursuing that.

SCHOLTZ: You couldn't get into the details of it as far as you would like? Bill, do you have a comment on that?

WILLIAM SANDER: There are two laboratories in the Army that address those issues. It's kind of not in the main stream of Paul's CECOM Ft. Monmouth's mission. But there's an office at the Harry Diamond Laboratory complex in Adelphi, Maryland

which is called the Survivability Management Office and all of you know Dr. Don Torierri who works there. So he would be a very good point of contact to discuss those issues. Then there is another laboratory called the Vulnerability Assessment Laboratory and they're not part of either of these organizations ... probably SMO, Survivability Management Office, is more akin to the communications kinds of survivability. The Vulnerability Assessment Laboratory is located at White Sands Missile range, it also does the communication electronics survivability. But they do a number of other things. They are spread out over a much wider mission area. But those things are a very strong concern and I suspect the Army is not doing a very good job of integrating those things early into the planning of systems.

SCHOLTZ: I may be in the wrong community to even hear about this stuff, so that could be the problem. Or maybe there really is something interesting there. John Treichler?

TREICHLER: I wanted to go back to, ... this will all come out at the end, but I did want to make one comment in response to something Chuck said. I sort of expected Bart Rice to jump up and make some comments too. I'm going to sound horribly like a systems person and an applications person rather than an algorithm person even though I think of myself more as the latter. You talked about features versus measures versus other things. I think this is my third non-answer. There's no right answer. You go to each particular customer who has a particular problem and you ask him what he cares about. And his answer might be, I want to know every signal from every other signal in the whole wide world. I'm going to put up a 15 ft. whip antenna and I want to know anything that impinges on it. Well, actually, relatively few people ask that question. They say, Hey, my particular problem is teleprinter signals and all I want to do is find those suckers and measure the parameters and then set up a demodulator and go, as an example. Or, I want to be able to tell that from a couple of other interferers or jammers. Or, I want to, I want to ... and every office you go to has a different view of the problem. And every time they tell you what the problem is you almost have to start at ground zero again, think about it and then look in your tool box and see what you've got. I tend to think about where the university people can contribute most is putting tools in the tool box and figuring out how good those tools are and what they are good for from an analytical/theoretical point of view so that when the more application-oriented or system engineering types come in and look at it, they can deal with some truths rather than some question marks.. And so I don't think there is a simple answer of measures and features. I think both of them are useful and it depends on the application, and there is no way to know ahead of time.

WEBER: Are you thinking that tools then are more important than robustness? I hear you saying that each problem is now addressed separately.

TREICHLER: No, I'm adopting the perverse point of view of, you tell me you've got a technique for measuring the feature or performing measurements, and I can almost always find an environment in which it's not robust. So I think even the concept of robustness is really environment-dependent as well. And you sort of have to do value-free judgments of these various techniques and figure out what each of them is good for independently. And it's only in the context of a

system design that you can really say what works, what's good, what's bad, what's robust.

SCHOLTZ: By the way, there's .... I'll get to Bart in just a second, I think he feels he now has to speak now that his name has been called. A couple of people mentioned to me they were really waiting for a break. I'd like to run this morning's meeting on the fly so if you want to take a coffee break for your own personal reasons, please go ahead. Some people have to check out, some people have to leave early and I don't want to lose any conversation during the break which has been happening the last few days. So let's keep it going.

WEBER: Bart, we knew we were going to get to you sooner or later ....

BART RICE: The ideas of signal classification and identification are slightly different. For signal identification you're looking for very specific things. For signal classfication you are trying to take in some energy and decide what kind it it is. Usually, the identification follows the classification. But, one area relevant to both that we didn't touch on very much is artificial intelligence. I think that AI has a role here for University research in this area, taking into account collateral information as well as signal data and making classfication decisions. Information about locations or other kinds of information about the signal that is not related to the waveform itself can be captured in rules which can aid in making decisions. One way to do that is take a Bayesian approach and adjust the a priori probabilities of certain signal classes in accordance with the collateral information. Development of a framework for doing that and, perhaps, development of an artificial intelligence system or data fusion system would, I think, make a valuable and interesting applied research project. Another possibility is one that you mentioned, Chuck, about sequential processing. Group three facsimile signals can appear over voice grade channels. There's a preamble for digital fax. If you only look at the modulation (after the preamble), it looks like many other communication signals. But, the preamble is generally a different modulation than the data. So, if you catch that preamble, and then a little later you catch the other modulation type, then that's a very strong indication that the signal is a facsimile signal. And, the decision was not based on any short little interval in which you say, "That's fax." Rather, it was based on a time history of your short-term classification decisions. This "sequential identification" idea has a rule-based flavor to it, and I think it can form a nice line of inquiry for University research. There are a number of others, too, in this same general area of applications of AI to signal classification.

SCHOLTZ: I was wondering whether the guy on your ... over on your left wanted to say anything for the record. [LAUGHTER] Is that a no? Samir, would you like to say something?

SAMIR SOLIMAN: I think among the set of questions you've posed, question numbers 2 and 6 are highly related. The receiver structure depends on which approach you'll use: the maximum-likelihood approach or the clustering pattern recognition approach. In other words, is it a parameter-measurement approach or feature-extraction approach? I believe if you use a maximum-likelihood approach, this will lead to an estimator-correlator type of structure which is nothing more than parameter measurement. If you use pattern recognition or clustering, this is nothing more than defining a set of features and then base your clustering technique on

that set of features.

WEBER: It could very well be. I wanted to make one more comment about Mark's presentaton. Coming to some extent from a radar background, I appreciate very much what he was doing, and I did detect essentially a gap between his research and the other talks. Maybe that is another way to express a gap that needs to be bridged, i.e. get his material more connected into the modulation characterization environment. Paul?

SASS: I just wanted to raise a question that hits us from time to time regarding modulation. From time to time our commanding generals get visited by people like AT&T who demonstrate that they are going to a higher order of modulations. We get more and more bits per Hertz offered to us on a platter, and the question that keeps coming up is how much in the way of spectrum occupancy can we really improve by getting higher order modulations in a non-ideal channel environment like the military has to deal with. We don't have a commercial telephone environment in our multi-channel systems. Does it make any sense to strive for more and more bits per Hz to solve the spectrum occupancy problem, or are we just wasting our time going to higher and higher order of modulation alphabets?

SCHOLTZ: I guess we could go back to Shannon and start from there.

WEBER: My first reaction is, you need to have knowledge of the channel.

SCHOLTZ: John Treichler has a comment I think.

TREICHER: Let me say this about that somewhat elliptically. AT&T's idea of a good propagation channel and what you can really get away with are very different. There do exist equalization techniques, methodologies which allow you to work in much worse situations, with much poorly sighted repeaters and so forth than AT&T, MCI and those people typically require. There are government agencies doing that sort of work and if the Army is interested in knowing about that, I'm sure there are methods to find out. So it really flips back the other way, do you really need it? You have to decide if spectral occupancy is important enough to go and make use of those technologies because they do exist. Basically what it comes down to is, if you've got enough SNR, there are methods to deal with the multipath problem, and, to some degree, the interference problem that you would find in your sort of environment.

SCHOLTZ: Any further comments for Chuck's summary? Was that a comment Ray? [LAUGHTER]

HALL: I collected some comments, some that Dr. Gardner had related to the continuation of that line to signal restoration work, primarily the inclusion of the multiple features, multiple optimization criteria type algorithms into their structures, a combination of both spatial and spectral methods into the same thing. I think the need to start addressing what the computational requirements are, when Brian Agee shows these nice scope-photolike things where signals are all nicely separated out, well, how long did that run and how long did that take and is it possible to put that in a finite box? What's it's going to cost Paul if he wants to do that? The second one, we've touched on that a couple of times already, and that is, Seymour Stein's comment of the .... Mark Wickert started out from the interceptor attributing to the communicator infinite power to disguise his signal, and then the same thing goes the other way, contributing to the interceptor infinite processing power from which to separate out interference or not. So I don't know if this

falls in the University's lap or maybe falls in some of the veteran analysis people that have been around a long time, to formulate and integrate a framework where we can trade-off like a radar equation, a simple thing, you've got to get so many bits across to the other side and if you go and try to make your amplitudes look Gaussian where you give up 5 dB. It's the comment that Dr. Scholtz had earlier, that's all folded into there. How much can you give up versus which trick, and how much are you willing to defeat the interceptor or the jammer? And, there's levels to that. You really need to define how to communicate. First the communicator needs to achieve some definable performance, some BER with some channel characteristic, then the first thing he would like to defeat is detection. So he'd like to be LPI in some sense. And the second level he'd like to then defeat would be interception and copy. So how much work does he have to do. Maybe he doesn't care that somebody knows where he is or knows that he's up and doesn't know where he is, that's another one too. You'd like to defeat the ability to be located by somebody. So, giving that up, maybe all you want to be is secure so you don't want somebody to be able to copy what your text is. That's either an encryption thing or a waveform isolation thing. Of course the last one is you just want to be able to talk. You don't care if someone gets your encrypted signal, maybe you're satisfied with that but you'd like to be jam proof in some situations. So I think, and obviously from the discussion earlier, other people think that there needs to be just a common set of system level analysis done there that everyone can work in and get the field level.

Point 3 is on the general area of equalization. It looks to me like the research is going for faster and faster convergent methods. They are getting pretty fast, 10 symbols and they are equalized. Dr. Proakis' work was very good because it showed the complexity in how many operations per symbol you've got to do to achieve that. It is kind of the environment that TRW is in. We'd like to see more work and it passes on that the Government would like to see more work done, reduced complexity algorithm. When you're up that 100 MBPS you don't want to have to do a 1,000 operations per symbol to equalize that channel, so the chart that Dr. Proakis showed something between the fast RLS and LMS. I think there's a big gap there in complexity and performance that needs to be filled and addressed. Then, in my session and of course in Dr. Peile's session, it's clear that waveform and information encoding and equalization are all tied right together. There's just no doubt about it. DFEs and trellis decoders, somehow they are the same. And that really needs to be addressed more from an integrated point of view and get the signal processing people, the information theorists and the comm theory people all together in one framework where they can address the same problem. It's a hybrid of those three fields.

Let's see, No. 5, this was a continuation of the one I addressed earlier, just a level field of comparison for the existing algorithms, more work like what Dr. Proakis did. That was wonderful, me as an industry user, I thought that was great. Now I can see the tradeoff. Those are the kind of tools we need developed to see if we should even launch in to look at adaptive lattice structures. The stability is nice but the computation is for 100 MB length, it probably eliminates that as a choice. But it's good to know that.

Then, probably on the equalization things, we need to address better implementation.

I know this work has been going on. It really didn't come out in our session but VLSI techniques, people have looked at microprocessor implementations and what not. The other one is the convergence rate versus the environment. I know John Cioffi has done some work in multipath environment. Really trying to define the multipath environment to where you can say, I need recursive least squares or LMS will do. How to find the dynamic environment such that a certain algorithm will address them? The other thing we talked about in the signal restoration area, or in the modulation characterization field, is the op's concept, or the operator issues. Are we going to have Brian Agee out there in a half-track looking at the screen, tweaking his multiple option criteria to separate these two signals or I guess Bart Rice's thrust is to make it all automatic so we can have something presented. I guess a little of that was addressed but I see that when you go to deliver equipment, that also needs to be addressed.

SCHOLTZ: Any comments on that? Yes Seymour, I'm glad to hear you speaking up.

STEIN: Some of the discussions on equalizers during the last two days reminded me of some work we did several years ago, and some lines of investigation that I regard as currently on the shelf. One recurring problem in equalizer design is the speed of convergence in trying to keep up with changes, which really becomes a problem when one of the terminals is moving, especially when there is multipath. Equalizers now are all designed based on the notion that the channel is momentarily stationary, in a fixed condition. The equalizer is a filter designed to react to that stationary condition, and then you squirm when you find out that the stationary condition doesn't last very long. The work we did several years ago involved a rather simple problem with a single delay path but with a moving terminal. We observed in that case that if instead of taking the common approach, we attempted to estimate the parameters of the multipath, namely the relative delay and relative Doppler, we could then construct an equalizer with time-varying weights which were good for as long as you could assume that the Doppler itself had not changed. That's a much longer time between updates. And I always am struck by the idea that as processing improves, maybe someone should re-examine that notion of trying to estimate the channel as part of the equalizer control algorithm rather than assuming that you don't know it; usually it's a set of discrete paths that are changing with some temporarily constant Doppler shift.

SCHOLTZ: I think John Treichler also had a comment. Was it the same one?

HALL: You know Seymour, the phone companies spend a lot of time trying to characterize their channels that way also.

TREICHLER: I always agree with everything Seymour says. I'm certainly not going to take issue with that. Actually I agree that a lot of the equalizer work that has been done today has been sort of, either nobody went out and did the channel measurements ahead of time or you weren't prepared to believe that what you measured was a general enough view of the world and people tend to drop back and say, well just use tap-delay lines again and hope for the best, but that's only to ratify his comment. But one comment I want to make, before I give the microphone back to him because I see him rising up in response, is that Dennis and I talked a few minutes ago and the point I would make is regarding level No. 5 of this level field comparison of existing algorithms. We have a large number of algorithms suggested, all

with their various proponents. There's not been time, money or energy to go out and look at all these things carefully and decide what each one of them is good for and bad for. And I think that needs to be done and that's certainly something suitable to the University environment. Carrying that a bit further, one of the real hard parts about that right now is that a lot of these algorithms are highly non-linear and we still think in the linear world, every time we turn around, all of our analytical techniques are essentially linear. New analytical tools are needed in order to get any sort of reasonable, believable analysis of some of these fancier schemes.

STEIN: Both Dennis and John's reaction made me think I didn't make my point. I was talking about on-line channel estimation, not off-line estimation of the channels, but real dynamic modeling, call it LMS modeling which has been done in some cases already.

SCHOLTZ: I think John Cioffi here has a comment.

CIOFFI: I don't want to disagree with anything that's been said. I do agree with all that but I think one of the directions that's wide open in equalization and no one has really touched on is on the interface of the link, network, and physical layers of use of the OSI model. A lot of the work discussed at this workshop is point-to-point type communications rather than point-to-multipoint or multiple access as well. I think if you look at it from a network perspective in the overall thoroughput that you're getting amongst many users, I believe equalization can help in those areas. And the only thing you tend to see along those lines right now are maybe spread spectrum approaches and more in the coding domain. I believe there is a merging of equalization technology and filtering, along with coding and protocols that is wide open. There are some fundamental problems in equalization of many users and how do you reject cross-talkers and such that could augment the existing technologies in that area. It's a hard problem, and fundamentally, I don't see any good solutions to that at this point.

SCHOLTZ: Any other comments on this particular topic?

STEIN: Not on this topic, but I've got something to throw out on another area. I don't know where it fits in. It regards tactical applications in electronic warfare. A couple of years ago we took a look at techniques for intercept, if you wish, of spread spectrum systems. As a part of that effort, we examined some illustrative AJ systems that we knew to be in the field. When you think from an algorithm point of view about spread spectrum intercept, you always try to devise algorithms that will work down into very low signal-to-noise ratios because you say, obviously, I'm going to have that problem. To our horror, we found out that all the battlefield tactical systems always operate at maximum power. And, in fact, the conclusion was that any almost casual intercept techniques that we knew of would work just Jim Dandy as long as we were within the radio horizon. The problem is a very real one and there are really two parts to it. Right now, we seem to take the attitude that if you're worried about jamming, we have to operate in the anti-jam, maximum power mode at all times because we don't know how we'll get to that mode after the onset of jamming. I'm not sure there are any easy alternatives but one bad aspect of that kind of operation, and I'm sure we're going to see it as more and more of these hoppers go into operation, is that the self-interference we cause ourselves may make us our own worst enemies. The other part of

that coin is, everybody talks about an enemy capable of doing everything. Everybody makes the assumption that jamming will simply be there, while the fact is that most of the jammers that are designed still do not engage in jamming Willy-Nilly; they respond only to a signal that they can hear. That notion that you have to hear the signal means that, despite the way we design our systems now, there is really a tremendous benefit to having LPI, meaning you try to operate at minimum power. The weird thing is that in the past, there have been systems that have started out being called LPI systems and before they ever got developed, the powers that be had changed them into AJ systems. They were no longer even being labelled LPI. I keep feeling that's a mistake, but the critical issue in all this is how can you do the coordination when you think of a fluid environment, so that you will somehow be able to convert over from one mode to the other. I'm not sure I am able to identify the nature of the research that's needed, but research leading up to tests that would prove the ability to do that sort of thing would go a long way towards changing the way the designs are set.

SCHOLTZ: Yes, Gaylord?

HUTH: I always had a problem when I worked on AJ and LPI systems whether anyone really wants LPI. I have worked on systems where LPI was really needed but when the system was used most users turned up the power to get AJ. This situation caused the people who needed LPI because of their vulnerability to be afraid to communicate and have their location determined. As I mentioned to Paul, I really do not know whether there is a conscious effort to be truly covert in terms of LPI or any other aspects. Besides the Army environment, there are situations such as report-back links where it is a burden to communicate for the field element that is required to report back. If the system is being jammed so the field element has to raise its power to report back, then it becomes more vulnerable to detection. In these situations, communications makes the field element vulnerable and if it becomes too vulnerable, there will be no report back.

SCHOLTZ: You know, we are drifting a little bit from the subject. I really would like to have this conversation on record. But we can finish Rob's panel and then I'll open it up to more comments. I think we have some more comments here too on protocols and variable rates, variable power level kinds of things. Maybe we will have something to focus on. Let's get Rob into this.

PEILE: I would first like to formally thank my panel for their contributions. I would also like to thank the Army Research Office for funding the research behind my talk. The summary of my panel is my own; my panel members are free to disagree.

The first subject we studied overlapped with the equalization panel, i.e. the subject of co-designed coding, modulation and equalization. My first comment is that, as a research area, it is going great. If every research area was making such rapid progress we'd be very happy. This is almost an inspirational area. I don't know if this is the greatest commercial success coding has ever had. I don't think so, probably a fire chip in an IBM computer in the 60s made more money. In terms of translating what is coding theory (and pretty abstract at that) into selling products, this is great. I'm not saying we should stop funding this area. I'm just saying the direction seems on-line. I have a mild caveat about the subject being driven by the telephone channel. We ought to remember that some channels do not have the

same delay constraint as the telephone channel. There is another research area that takes these new techniques and considers how they can be extended when there isn't a delay constraint. You can only do better. I see this as a future area. There are techniques being developed that do have relatively high delay throughputs. I know Vedat has done some work in interleaving techniques and I've done some myself. John Cioffi brought up the point that there is work to be done in point-to-multipoint techniques. That was a good point.

My next point was first brought up by Paul Sass. How do we transfer these magnificent techniques (which are getting close to Shannon's capacity for the bandlimited channel), make them robust, and bring them down to the power-limited channels that the military communication environment are more traditionally interested in? There's plenty of machinery around which could potentially do this. I wouldn't say it's being done at the moment. The benefit is more bits per Hz under some harsher environments. I'm sure that must be of research interest.

Alan Levesque surveyed hybrid ARQ techniques. It's my belief that we need a better global view. It's also my belief that if I send everybody off into a set corner of the room with a bit of paper, everyone could design a decent type-one or type-two hybrid scheme, and subsequently we'd all argue for years about which one was best. Recently, I have seen several code combining, type-one and type-two ARQ times appearing, all with impressive performance graphs. We might understand these schemes in isolation but I don't think we understand them globally. As I said yesterday, the system issues are tricky. You can analyze the components in isolation and get nice papers out of them but when you try to build a system around them, it's not clear how they all fit together. I see that as a research area where a lot more work needs to be done. For example, code-combining is getting on in age, but we still can't analyze it completely. I am not sure that it's the best thing to do without modification.

The last point (what most of the debate was about) is adaptive FEC as an adaptive system component. The debate (more or less) centered around HF except for some mention of the meteor channel. How can adaptive FEC serve as an adaptive system component? I'm reminded at this point of a phrase that an engineer came out with years ago. He was bewailing his problems and a director decided to be supportive and said, "You haven't got problems, you've got opportunities." The engineer replied, "I am surrounded by insurmountable opportunities." That's a pessimistic view. I think research has to be done on a two-way basis. Explore the possible use of decoder error locations as a source of real time system assessment. As I said, if the information is there, why throw it away? It might be difficult, but why throw it away? The research's onus is to expose the possibilities to the systems people. There's only a limited amount you're going to achieve in adaptive FEC on a link-to-link basis.

A lot of the objections to adaptive coding reflected a basic premise. I have never met an error correction application, a successful one anyway, which was really about error correction. It was about something like data density, or about lower power, or about some other parameter. The real impact of coding, in the successful applications, has been to alter a classical tradeoff. I think that adaptive FEC researchers have to look at how the classical tradeoffs are altered by the presence of coding. That means they have to know what

ADAPTIVE CODING - RESEARCH DIRECTIONS

1. Coding/Mod^n/Equal^n for BL Channels

A) Going Great!

B) Remember some Channels do not have delay constraints.

C) Transfer techniques down to PL/BL Channels.

2. Hybrid ARQ/FEC Schemes

A) Need better global view of Code-combining/Type 1/Type II combinations + performance

B) System Issues

3. Adaptive FEC as an Adaptive System Component

A) Surrounded by Insurmountable Opportunities!

B) Explore possible use of decoder error locations as source of real-time system assessment. Expose possibilities in Systems/Networks.

C) Examine how the Classical trade-offs are altered by presence of coding.

D) LISTEN TO OBJECTIONS: SEARCH FOR DATA + EXPERIMENTS

E) Trial Systems

the classical tradeoffs are. And that means they've got to listen to people who have been working in adaptive systems. (Which is the reason why I insisted that Seymour was on the panel.) You have to know the classical tradeoffs and study how they are altered by the presence of coding.

I think adaptive coding has to learn to listen to objections. The objections are well founded in many cases. You have to search for channel data and experiment in some real channels, which was a point made by Bob Scholtz. This is a major problem. How can coding people in an academic environment have access to communication channels? This problem is particularly acute if you're going to crucify a coding person for not having analyzed real channels when he gets up and presents his theoretical analysis. I am constantly on the scrounge for data. Now I've heard other people say this, and I sometimes suspect it's crocodile tears, and they really want to have Gaussian noise analysis because it's a lot easier, and the sums are easier. But yes, I really do mean it. You need channel data, especially for HF which I regard as one of the ultimate challenges of error-correction. Ultimately, I don't think anyone will believe adaptive coding until you have some trial systems and some experiments and you can run them. And I think research has to be guided towards that line. Let's do some theoretical work. Let's look at the classical tradeoffs. Let's position ourselves to the point that people believe us enough that they are willing to consider an experiment that's worth doing. That is a research direction which I believe is important.

O.K., that's a very personal summary of research directions in adaptive coding. It's a formative area so there's a lot more personality in it, perhaps more than the other panels.

SCHOLTZ: Any questions or comments directed at Rob's summary? Are there any adaptive coders?

HERMAN BUSTAMANTE: I mentioned earlier that we at Stanford Telecomm have put together a system which in fact might provide some information relating to the kinds of systems that Rob Peile was just commenting on. If I might have a minute, I have prepared some slides, which describe the system I'd like to tell you about. Could I show you what I have? Thankyou.

What we are building at Stanford Telecommunications is a system we call the 1105B communication sub-system. It operates in satellite channels and primarily it is a TDMA system. The features that might be of interest to you [SLIDE 1] are shown here. It can operate either in single frequency or multi frequency mode. It includes not just what you call the modem function but of course a multiplex function. It can accommodate up to 500 baseband channels and contains an automatic patch panel capability. In other words, you can put in or take out any channels automatically given that you're providing the baseband equipment.

We are incorporating automatic power control and we are also doing adaptive, if you will, modulation and coding characteristic control.

What we have is a problem that has been mentioned by several speakers and particularly by Rob Peile, and that is that we have both a bandwidth constraint and a power constraint that we have to stay within. In other words we are operating through a satellite, and the satellite resources are partitioned into bandwidth portions and power portions, and you are not permitted to exceed these.

Let's go back to the TDMA system. We are operating as a network with one user op-

erating as a master, let's call him our control terminal, and a number of field terminals. We can have up to 32 field terminals. The problem is what do you do when you have rain? Well you can say, let's just make sure we have adequate power and burn our way through. Well that's nice but it might not be the most cost effective overall way to do things. What we have implemented is 3 methods of control. One of them is to control the uplink power but if you look at the link operating from the master to the field, uplink power doesn't do you any good on the downlink. It will help you on the uplink obviously but on the downlink you want to do other things. One of the things you can do is vary the effective data rate transmission. In other words, we vary the effective data rate by extending the duration of the bit and in that way we can provide up to 12 dB additional margin in the link or we can provide 12 dB compensation for downlink attenuation. The problem we ran into was that we didn't like the performance we were getting because the control steps were very coarse. As you go from a symbol length of 1 to 2 to 3 to 4, you're going in increments of 0 dB to 3 dB to 4.7 to 6. The jumps are quite large and we wanted to have a finer control. What we decided to do is to use coding rate variation in order to provide a finer control capability. What we've implemented is a system that can compensate up to 12 dB rain-fade attenuation in 1 dB steps. It can accommodate rain fades with up to 6 dB per minute fade rate and it can do it quite nicely for the system as a whole. The numbers I'll show you may not seem very impressive but if you look at it on a system basis, I think you will agree that it can be quite useful.

This summarizes [SLIDE 4] the parameters that we've ended up with. For the uplink power control I believe the right number is 2/10ths of a dB, I don't remember exactly. That's the control we have on the transmit level. We can accommodate at least a 20 dB range and we can select the limits. For the data rate variability, as I mentioned, we go in increments from 1 to 2 to 3 up to 16 which gives us 12 dB. And we are providing coding rates of 1/2, 4/7, and 3/4. This enables us to provide gain steps that are on the order of 1 dB over the entire range. And so here we have a combination of power control, coding rate and, if you will, the effective data rate to give us the total control capability I was talking about. The reason I wanted to come here is that I'm not satisfied with the overall performance in terms of how quickly we can respond, and in terms of how fine we do things. I'd like to be able to provide a more rapid capability so that we can actually track these rain fades much more rapidly and do the overall compensation better. That's the kind of help I'd like to find from whoever would like to tell us how best to do it.

We are also looking at other types of links, links which are not on continuously but which sort of hop but I can't go into the details. The system I have in mind may want to communicate with one source, then another, then another and I need to have very rapid acquisition, very rapid characterization of the channel so I can establish the parameters I want to operate with. That's why I was very interested in coming here. So I'm hoping that with a combination of the methods that we've used here which is to do signal quality estimation on both signals from the field and the master terminal and combining that with equalization, adaptive coding, what have you, I'd like to think that the next generation of equipments will be even finer control, faster acting and maintain performance within fractions of a dB rather than 1 dB or so. That's

1105B COMMUNICATION SUBSYSTEM

FOR

SATELLITE CHANNELS

HERMAN BUSTAMANTE

1105B COMMUNICATION SUBSYSTEM

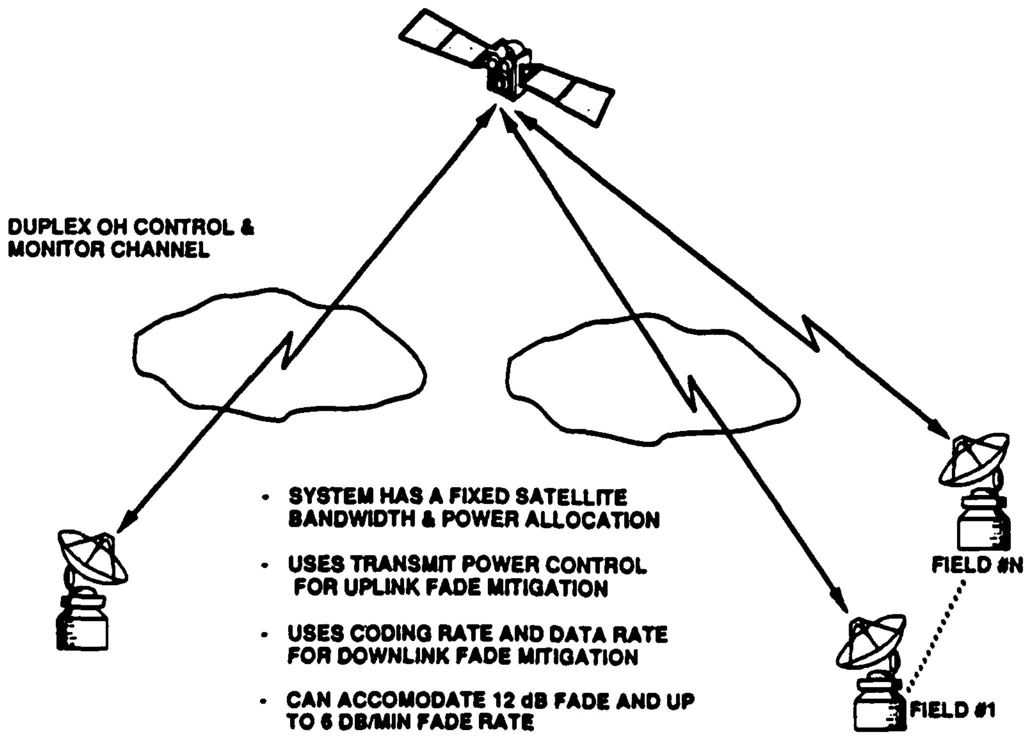
- ACCESS MODE	TDMA
- MODULATION	BPSK, QPSK
- CODING	CONVOLUTIONAL CODING VITERBI DECODING K = 6, 7; R = 1/2, 4/7, 2/3, 3/4
- BURST RATE	64 Kbps TO 12.288 Mbps
- BASEBAND RATE	300 bps TO 2.048 Mbps, SYNC 75 bps TO 19.2 Kbps, ASYNC
- # OF BASEBAND CHANNELS	UP TO 500
- # OF NODES IN SUBNETWORK	UP TO 32

THIS IS AN *ENHANCEMENT* FOR AN EXISITNG SYSTEM WHICH HAS APPROXIMATELY 150 OPERATING MODEMS IN THE FIELD.

STANFORD TELECOM

SLIDE #1

TSO 5134 6/1/89 LT



- SYSTEM HAS A FIXED SATELLITE BANDWIDTH & POWER ALLOCATION
- USES TRANSMIT POWER CONTROL FOR UPLINK FADE MITIGATION
- USES CODING RATE AND DATA RATE FOR DOWNLINK FADE MITIGATION
- CAN ACCOMODATE 12 dB FADE AND UP TO 6 DB/MIN FADE RATE

STANFORD TELECOM

SLIDE #2

TSO 6/2/89-36740

1105B COMMUNICATION SUBSYSTEM

- TDMA
  SINGLE FREQUENCY
  MULTI FREQUENCY
- MULTIPLEXER
- AUTOMATIC PATCH PANEL
- AUTOMATIC POWER CONTROL
- ADAPTIVE MODULATION/CODING CHARACTERISTICS

SLIDE #3

STANFORD
TELECOM

TSO 6125 6/1/88 LT

POWER CONTROL CHARACTERISTICS

- UPLINK POWER CONTROL RANGE
  < 0.5 dB STEPS
  > 20 dB RANGE, LIMITS SELECTABLE
- DATA RATE VARIABLE AS INTEGERS OF SYMBOL DURATION
  SYMBOL PERIOD = 1, 2, 3, 4, ...16
  BURST RATE = R, R/2, R/3, R/4, ... R/16
  GAIN STEPS = 0, 3, 4.7, 6, ... 12 dB
  TOO COARSE FOR LOWER VALUES
- GAIN STEPS DESIRED ≤ 1 dB
- COMBINE DATA RATE CONTROL WITH CODING RATE CONTROL
  CAN MAINTAIN OPERATION WITHIN < 1 dB FOR FADE RATES OF 2 dB/MIN
  CAN MAINTAIN OPERATION WITHIN < 2.5 dB FOR FADE RATES OF 6 dB/MIN
  MAX FADE DEPTH OF 12 dB

SLIDE #4

STANFORD
TELECOM

TSO 6136 6/1/88 LT

all I wanted to tell you about.

SCHOLTZ: It looks like we have some questions. Dennis?

HALL: What kind of ... right now you're tracking 6 dB per minute in rain fade?

BUSTAMANTE: The reason we came up with that number is that I tried to find in the literature just how fast we needed to track things when operating at X-band. I've done a lot of measurements at KU-band but I don't have personal experience at X-band. Our people on the East Coast, however, have done some studies that indicate 6 dB/min to be fast enough to accommodate a very high percentage of what you would run into with X-band. That's how we chose that number.

HALL: What's your goal though? You said you'd like something that adapts faster?

BUSTAMANTE: Well, what I meant by faster is that we don't follow the signal precisely. What we have done is to define two ranges of operation. Rain fades up to 2 dB/min and rain fades up to 6 dB/min. Rain fades up to 2 dB/min we can track within 1 dB. Rain fades of 6 dB/min we can only track with 2 1/2 dB. It's not a simple problem. When you stop to think about it you've got power controls going on, you're worried about your AGC, the degradation of your Viterbi decoding on and on. There's a whole multitude of things that come into it, so 2 1/2 dB looked like it was a realistic goal which we have pretty well been able to achieve. We can do better in terms of a n-gative margin, if you will. we are never farther away than 2 dB in a negative sense. We do sometimes overshoot in the positive sense which means we are providing two and a half dB more power than we really need. But we decided that was something we could live with because we don't expect all terminals to be suffering rain simultaneously. Besides which, this condition only comes into being during the ramping up or ramping down phases. So looking at it that way, it seems like we could tolerate a couple of terminals going through this kind of condition and still not be concerned about overdriving the satellite if you will, and causing any excess power utilization. It seemed like a very safe assumption in that case. Now when I say I would like to do it more rapidly, I'm happy being able to track 6 dB fades per minute but I would like to be able to do it better than two and a half dB, that's what I meant, finer control, more rapid recognition.

PEILE: I have a question. If I could remember right, it's a true statement that your Viterbi decoder gives you other statistics or some quality outputs which at the moment you're not using.

BUSTAMANTE: It's probably true but we sort of did it with what we had available at that point in time. You know better than anyone else what we've been doing.

PEILE: Yes, because I designed the chip. I ought to say by the way that I worked for Herman as a consultant and I found out more about the system right now than I did by going to Stanford Telecommunications. [LAUGHTER]

BUSTAMANTE: We had a hard enough time understanding what we wanted to do with a Viterbi decoder let alone solve all the other systems. If I told you about this, we'd still be working on it.

SCHOLTZ: I think Paul Sass has a comment here.

SASS: I just have a general comment on the transfer of all this great adaptive coding work into real systems. I've had to lose a lot of sleep over the problem of where in the system you integrate things like very sophisticated coding strategies. Very often we are not given the freedom of designing a whole

new radio. What we're given is a radio that provides an uncorrected 16 Kb data port, for example, and we are asked how can we best improve the error performance of this with a particular application. What I'm asking you to consider is the impact of a security function. I know this community doesn't like to worry about security but time after time we get bitten by the fact that there is a security function that is effectively separating the channel and your observation of the channel from the coding that you do. And if you view a security module as having an effect on the channel error statistics, I think you have to change some of your strategies.

PEILE: I agree totally. Most coding discussions I get into seem to split schitzophrenically between which side of the crypto you're on.

HUTH: When you are talking about where the crypto device is, are you talking about having the crypto before or after the decoder?

SASS: I'm talking about having the crypto between the decoder and the front end. It depends on whether you're transmitting or receiving for it to be called before or after.

HUTH: The systems I have worked on usually have the channel coder before the crypto device. The channel is in the middle with the error correcting code closest to the channel. The crypto devices are further from the channel yet. You may have error detection coding beyond the crypto device. Since the crypto device is away from the channel, you still have channel measurements. Even though you do not have the data that you can use, you do have the channel measurements out of the decoders. Therefore, you can still do these kinds of adaptive decoding schemes even though you do not know what the symbols are themselves.

SASS: Again, you're assuming there that you had the channel coder on the channel side. That assumes you had the freedom to design it in from the beginning. I'm saying very often we're given a radio that doesn't have the error control we want. We are just given a data port through which we ....

HUTH: So you want to add a decoder applique to a system that already has a crypto device between the channel and the applique.

SCHOLTZ: Yes, you're forced to that.

HUTH: O.K., I understand. I have comments on that.

SCHOLTZ: Any other comments on Rob's presentation?

PEILE: Any comments I missed? Any volunteers of data?

I would like one comment on Gaylord's point about LPI. It might be semantic but when I was listening to the intercept receiver people it struck me that low probability of intercept was not necessarily low power. It may be low probability of detection is a different thing than intercept. I think you can make high power signals that look like noise to the intercept receivers that we've heard discussed. I discussed the idea of rotating a multi-dimensional constellation by just a transformation. That just seems to make it very hard to classify in the present two-dimensional analyses.

HUTH: When I say LPI, I really mean covert. I am really talking about energy that is being transmitted that could be used for location rather than determining the transmitted message.

PEILE: Low probability of detection.

HUTH: Yes, *low probability of detection* (LPD) rather than necessarily intercept. I know of systems that need LPD. It is not clear from Paul's presentation whether LPD is important to him or not. The question is what

systems are important for general study.

SCHOLTZ: I think Bill Lindsey had a comment.

LINDSEY: This comment is strictly directed to Herman. I found his system description quite interesting from the terminal point of view, but what satellite do you have in mind to play that waveform design? Is there a processing satellite you have in mind? [RESPONSE WAS NOT RECORDED.] It's a bent pipe? O.K. I didn't get that part.

SCHOLTZ: Ray ....

RAYMOND PICKHOLTZ: I'd just just like to make a summary observation on this whole meeting that we had here. It's true that the title of this meeting was Advanced Communications Processing, and it seems to me that virtually everything was focussed on what people in networks call the level I, physical layer problem. But as Paul Sass pointed out, and John Cioffi, I think stole a little bit of my thunder, there are the network issues. Some of these issues and problems migrate up higher into the layering structure. In fact if you talk to people who've built commercial networks, they'll tell you that the major cost and the major problems are in things having to do with software issues rather than hardware issues. There are a great many unsolved, in fact, a great many undefined problems and if you extend the notion of advanced communications processing to include those things, that's something that I think, in the next decade, will dominate the issues. Very often at the lower level when we talk about coding and signal processing, we are very pleased if we get 3 or 5 dB improvement, but there are probably 10, 15 and 20 dB effects out there when you get into networking and there are a lot of open theoretical problems in routing and flow control and in protocol verification. Perhaps also, these protocols could be used for signal exploitation. I think that was mentioned by Bart. Perhaps it might be a good idea in a future meeting like this to get the level I community, which most of us here are, more aware of some of these higher level problems so that we can address them with the kinds of background and facilities we have. Perhaps we could even exploit these ideas to further develop adaptive processing.

One final comment is that, at least in the commercial world where fiber optics is beginning to dominate many of the transmission problems, many of the things I've heard talked about aren't even an issue. That is in the traditional sense. Nevertheless the network issues remain an issue, at least at the third and fourth level. Fast packet switching which contain fascinating theoretical ideas ought to be looked at more closely in a future meeting like this. Thanks.

SCHOLTZ: I think that's great. Are there more comments on adaptive coding before we sort of wrap up this meeting?

PEILE: I would say that such a thing as adaptive coding in networks hasn't been touched in this meeting. I'd like to mention this as a known omission.

KENNETH WILSON: As long as we're throwing out possible research topics, one that I had stumbled across within the last two months is the issue of local communications in the community of space station. One proposal at present is to use a radio frequency network based on KU-band with omni-directional antennas from extra vehicular astronauts performing Extra Vehicular Activities (EVA). Because of public relations issues NASA wants to be able to bring back to us, high definition, full color, full motion video television broadcast into your living rooms so that they can maintain the public relations necessary to maintain funding. As

a consequence of that there is a requirement to be able to bring back from that EVA astronaut 67 MB per second signal to the space station and then for relay down to the earth. One of the things that people have realized is that we have a severe multipath problem when we're working from an omni-directional antenna to another one. There's no intention to track the astronaut from the space station side. So if you need a research topic, why not model the space station and come up with some adaptive equalization that will allow us to go ahead and implement that channel and succeed in this sense. And if it becomes impossible to do that, we'd like to know that too because I'm interested in throwing that RF link away and providing them an optical communication system.

SCHOLTZ: I see several people with hands up now.

HUTH: I just wanted to add something to that comment. It is not only trying to bring back a lot of data but the EVA is very power limited. The required power sources add weight which is a huge problem. Therefore, you really have a power limited problem. There is a very interesting tradeoff between power and allocated bandwidth. I think it is a great area to study.

SASS: Just a brief comment. You've described something that sounds very similar to the mobile command post problem that we have. If we assume that we're going up to omni-directional, very high frequencies where we don't want to track, it sounds very similar. It's something I want to hear about.

SCHOLTZ: Bill Lindsey ....

LINDSEY: I guess my reaction to that is, sure this problem is analogous to the Army's multipath problem. I don't think it's nearly so severe, however. But I really have to say that I'm not sure what the multipath looks like with regard to this problem and what you're saying is that a first step before you deal with RF communications would be to attempt to understand what the multipath phenomenon would be and not commit to it and go and have perhaps to fix it. I guess I would go along with that. But it seems to me that the multipath problem could be characterized rather easily, in fact some work has been done on that particular problem from that space station environment. I know that the light and optical communications would compete with the radio frequency k-band baseline, but there are issues also that perhaps aren't answered from the perspective of NASA having to do with utilization of light between moving terminals and things of this nature.

SCHOLTZ: Maybe we don't want to solve the problem here. We just want to get it out in front of people. Any more comments? I'd like to open this up to comments about this workshop. Maybe critique it, was it too wide, too narrow, what subjects would you like to see in a future workshop. I think Ray has already opened up one. Let's go from there. Are we close to what you're going to talk about?

TREICHLER: Close, but not exactly. Actually I wanted to revisit a contingent issue that came up yesterday that I've been biting my tongue for a day and just can't stand it any longer. By the way I agree with everything Ray said while we were in passing. We joke in our business about SMOS, a Small Matter Of Software. You know we're going to build something and it's a small matter of software to change it. In fact the software is a factor of 10 times the hardware required, thus standing on its ear every argument made about, Oh, let's just use off-the-shelf hardware, then we'll just write a piece of software for it. So you have to look at those issues as

well. The other thing is, we've been fighting here for a dB, half a dB, 2 dB, then I heard in another talk we threw away factor of 10 to 1 in data rate on the fiber optic LAN, because that's just what the protocol requires. So it's interesting how the rules change when the bandwidth is free.

My comment was about customers. That came up yesterday. I think the very fact that we're here is an indication that this particular customer has a fairly clear idea of what needs to happen in order to make progress in a scientific and technical field. I just wanted to tell you that I view customers in sort of three different grades and since nobody here that I know of is paying me, I think I don't have to be too careful. The customer I fear the most is the one who walks in and says, "I want to buy this. Here's my spec. I don't know where I got my spec but here it is. And what do you have off the shelf that can do this? And I really don't want to talk any technical details." That just sort of scares me to death. The next flavor of customer that is definitely a big notch up from that, obviously anyone who'll talk to you is better than one who won't. The next customer up that I like is the one who provides what I call the outrageous requirements. He's the one who says, "Not 2 dB better, not use the same radio and do this. But let's look out there a ways 10 years from now here's where I want to be. I want to be able to have every infantry man in the Army have something that weighs 2 oz., consumes no power and allows me to communicate back to the President if necessary. Because the President thinks he ought to dictate every action on the battlefield." But you say, you can't do that. Well, if you set an outrageous enough requirement, far enough out into the future, that makes you kick over stones, not get locked into the way I'm building it now, not get locked into TRITAC, this or that or whatever. You can start, in the industrial community, the academic community, looking out there and saying, "What research am I doing that generally leads towards that?" And the step up from there in terms of wonderfulness of customers is the one who'll set out that outrageous requirement many years into the future and puts together a funding program in University and industry and works with you. I'm not talking about outrageous amounts of money either, I'm just talking about enough to create a body of knowledge about the technologies required to achieve that objective. So then, 5 or 6 years from now when all of a sudden somebody from Congress looks up and says, "Oh my God, we need the following," there's a body of technology required to start solving the problem. Every little contract funded, looking at every little piece of technology, may not get used, but it was all necessary to create the people and the interest to be able to come to workshops like this and so forth, and share ideas. There's got to be a body of knowledge. And if you carry that one step further, the customer with the outrageous requirement who has some money and interest to nurture along different organizations, and then, in addition, has some test bed capability, where it's possible to come try your ideas out every so often, then you have reality to go along with all of your ideas. I have certainly met the ... "I need 10^{-6} bit error rate and I don't care anything about it" customer, but my personal choice is to find the others, to find the ones who are more technically interested, have technical judgment, in spite of low pay in the Government, have persevered and have been willing to stick in there for the grander good.

SCHOLTZ: More comments ... Gaylord

looks like he wants to say something.

HUTH: I want to comment on something that Paul said about the crypto devices. You can do adaptive decoding techniques like those Rob Peile was talking about in terms of identifying the channel including the crypto device. My question is whether that problem falls into the University research area. Maybe Rob can tell you all about how to determine the underlying process that is creating the errors through the crypto device. However, sometimes you have to know how the mechanisms of the crypto device work in order to actually determine the error process. I know you can use Rob's approach because the crypto device puts out certain error statistics that you can work with and exploit. My question is, what can be done at the University?

SCHOLTZ: I don't know where it can be done.

PEILE: The first thing to say is that this isn't a necessarily classified area. To know the effect of a classified device on error statistics might not mean that the statistics are classified. I've worked in coding for 5 years in the States and I've never had a clearance. I've worked in a lot of cryptographic environments and it hasn't stopped me. The definition of the channel can change. I would point out that soft decision information is practically a definition of a security breach if it gets on the red side of a crypto. Moreover, if you're going to be on the red side of the crypto, you might as well retreat even further into the OSI model. This ties in with Ray Pickholtz's last point.

SCHOLTZ: More comments, I've heard software development, I've heard questions about networks and the protocols and the vulnerabilities at that level. Are there any other things that people think really need some exploitation at a workshop? [PAUSE] You must have solved a lot of problems at this meeting! Does anyone have any executive comments?

Paul Sass raised a question. He thought people may have been subdued a little bit by the fact that they're speaking into microphones and these words will appear in print. This is an unclassified workshop and has always been. In theory that should be enough said one way or another. But I'd open the floor to the process of recording comments. I think Bill Sander and the Army feel it's a very fruitful thing to do, to really keep a record of this for other people who want to hear what we said. Does anyone feel subdued by just this process of recording? [PAUSE] They are all subdued, they don't want to be on the record to say it. [LAUGHTER] I'll turn the mike off and we can have this discussion. *[NO TAKERS] I wore my purple shirt* today because I thought Lloyd Welch's video tape needed a little more color! Are there any other comments about the workshop and its management? Bill [Sander], do you have any more comments as our sponsor? [MISSING RESPONSE]

I myself would like to thank the people who came to the Workshop and especially to the last session because it is a long drag to go through two and a half days in this environment and ignore going horseback riding off hours, on hours and that sort of thing. I'd also like to thank the people that came from slightly different areas because I'm a communication theorists and most of my cronies are too. It's nice to see people from the Signal Processing community and some of the Channel Measurement people and we really appreciate your participation. We hope some of our future topics will be of interest to you and you'll join us at a future workshop. Thank you all very much and let's call it off for today.

THE END

ATTENDEES - CSI-ARO Workshop
Ruidoso, New Mexico
May 14-17, 1989

AGEE, Mr. Brian .... (916) 661-1243 or 1102
AGI Engineering, 741 Daniels St., Woodland, CA 95695

AMES, Mr. Steve .... (415) 852-4129
Ford Aerospace & Comm. Corp., MS G79, 3939 Fabian Way, Palo Alto, CA 94303-4697

BELLO, Dr. Phillip A. .... (617) 271-8081 or 6262
MITRE Corporation. R312, Burlington Road, Bedford, MA 01730

BRADSHAW, Dr. C .... (703) 883-5667
MITRE Corporation MZ245, 7525 Colshire Drive, McLean, VA 22102-3481

BUSTAMANTE, Dr. Herman .... (408) 980-5633
Stanford Telecommunications, 2421 Mission College Blvd., Santa Clara, CA 95054

CASSELLS, Ms. Cathy .... (213) 743-8306
Communication Sciences Institute, University of Southern California, Department of Elec Eng., Los Angeles, CA 90089-0272

CIOFFI, Dr. John .... (415) 723-2150
Stanford University, Information Systems Lab., EE-Department, Stanford, CA 94305

COMPTON, Dr. Ted .... (614) 292-5048
Ohio State University, 2015 Neil Ave, Columbus, OH 43210

CREPEAU, Dr. Paul .... (202) 767-3397
Electronics Eng. Comm. Sci., Naval Research Laboratories, Code 5523, Washington, DC 20375

DIZER, Dr. Tom .... (703) 347-6415
The Director, USA CCSW, Attn: AMSEL-PD-SW-TPO, Vint Hill Farms Station, Warrenton, VA 22186

EYOBOGLU, Dr. M. V. .... (617) 364-2000 x7248
Manager Modem Research, Codex Corporation, 20 Cabot Blvd., Mansfield, MA 02048

FRIEDLANDER, Dr. Benjamin .... (415) 323-2608
Signal Processing Technology Ltd., 703 Coastland Drive, Palo Alto, CA 94303

GARDNER, Dr. William .... (707) 944-0648
Dept. of Elec. Eng. & Comp Science, Univ. of Calif Davis, Davis, CA 95616

GAULT, Dr. James .... (919) 549-0641
SLCRO-EL , Electronics Division , Army Research Office , P.O. Box 12211 , Research Triangle Park, NC 27709-2211

GHOSH, Ms. Monisha .... (213) 743-0907
Communication Sciences Institute, University of Southern California, Department of Elec Eng., Los Angeles, CA 90089-0272

HALL, Dr. Dennis .... (213) 812-1725
TRW, Mail Code R12/2688, One Space Park , Redondo Beach, CA 90278

HUANG, Dr. T. C. .... (213) 416-7182
The Aerospace Corporation, M6/210, P.O. Box 92957, Los Angeles, CA 90009

HUTH , Dr. Gaylord .... (213) 641-8600
AXIOMATIX Corporation , 9841 Airport Blvd. , Suite 1130, Los Angeles, CA 90045

KAWAGUCHI, Dr. Dean .... (213) 813-7894
TRW, Mail Code 03/2234, One Space Park , Redondo Beach, CA 90278

KUMAR, Dr. Vijay .... (213) 743-5387
Communication Sciences Institute, University of Southern California, Department of Elec Eng., Los Angeles, CA 90089-0272

LEVESQUE, Dr. Allen H. .... (617) 466-3729
GTE Govt. Systems Corp., EDCD - 2410, 100 First Ave., Waltham, MA 02254

LEYENDECKER, Dr. Robert L. .... (703) 347-6492/6493
U.S. Army, Center for Signals Warfare, AMSEL-RD-SW-TRS, Vint Hill Farms Station, Warrenton, VA 22186-5100

LINDSEY, Dr. William .... (213) 743-2349
Communication Sciences Institute, University of Southern California, Department of Elec Eng., Los Angeles, CA 90089-0272

LINSKY, Dr. Stuart .... (213) 813-8034
TRW, Mail Code 03/2240, One Space Park , Redondo Beach, CA 90278

MONTENEGRO, Ms. Milly .... (213) 743-5537
Communication Sciences Institute, University of Southern California, Department of Elec Eng., Los Angeles, CA 90089-0272

MULLIGAN, Dr. James J. .... (703) 347-6493
Director U.S. Army , Center for Signal Warfare , ATTN: AMSEL-RD-SW-PE , Vint Hill Farms Station , Warrenton, VA 22186-5100

PAWLOWSKI, Dr. Peter R. .... (213) 812-0170
TRW, Mail Code R12/2173, One Space Park, Redondo Beach, CA 90278

PEILE, Dr. Robert .... (213) 743-0060
Communication Sciences Institute, University of Southern California, Department of Elec Eng., Los Angeles, CA 90089-0272

PETERSON, Dr. Roger L. .... (602) 732-2452
Motorola Inc. , Strategic Electronics Div. , G 2242, 2501 S. Price Road , Chandler, AZ 85248-2899

PICKHOLTZ, Dr. Raymond .... (202) 994-6538
School of Eng. & Appl. Science, George Washington University, Phillips Hall, Washington, DC 20052

POLYDOROS, Dr. Andreas.. (213) 743-7257
Communication Sciences Institute, University of Southern California, Department of Elec Eng., Los Angeles, CA 90089-0272

PROAKIS, Dr. John.. (617) 437-4429
Northeastern University, 411 Dana Hall, 360 Huntington Ave., Boston, MA 02115

PURSLEY, Dr. Michael... (217) 333-2966
University of Illinois, Coordinated Science Lab., 1101 W. Springfield Ave, Urbana, IL 61801

RHODES, Dr. Stephen D.. (703) 347-6492/6493
U.S. Army, Center for Signals Warfare, AMSEL-RD-SW-TRS, Vint Hill Farms Station, Warrenton, VA 22186-5100

RICE, Dr. Bart.. (408) 756-4892
Consulting Engineer, Lockheed Missile & Space Co., 0/6231 - B/150, 1111 Lockheed Way, Sunnyvale, CA 94089-3504

ROLLINS, Mr. David... (213) 743-0907
Communication Sciences Institute, University of Southern California, Department of Elec Eng., Los Angeles, CA 90089-0272

RUDE, Mr. Michael .... (213) 743-3962
Signal & Image Processing Institute, University of Southern California, Department of Elec Eng., Los Angeles, CA 90089-0272

SADR, Dr. Ramin ....
Jet Propulsion Laboratories, 4800 Oak Grove Ave., MS 238-420, Pasadena, CA 91109

SANDER, Dr. William A.. (919) 549-0641
SLCRO-EL , Electronics Division , Army Research Office , P.O. Box 12211 , Research Triangle Park, NC 27709-2211

SASS, Dr. Paul .... (201) 532-0164
Director , U.S. Army Commun & Elec., AT:. AMSEL-RD-COM-AC-B , Center for Comm. Sys., Ft. Monmouth, NJ 07703-5202

SATORIUS, Dr. Edgar H.. (818) 354-3016
Jet Propulsion Laboratory, 4800 Oak Grove Ave., MS 238-420, Pasadena, CA 91109

SCHIFF, Mr. Maurice ....
Engineering Staff Specialist, General Dynamics Convair, P.O.B. 80847, San Diego, CA 92138

SCHNEIDER, Dr. Allan .... (703) 522-0046
Cybercom Corp., 4105 N. Fairfax Dr., Arlington, VA 22203

SCHOLTZ, Dr. Robert... (213) 743-5546
Communication Sciences Institute, University of Southern California, Department of Elec Eng., Los Angeles, CA 90089-0272

SIMON, Dr. Marvin .... (818) 354-3955
JPL , M.S. 161-228 , 4800 Oak Grove Ave. , Pasadena, CA 91109

SOLIMAN, Dr. Samir .... (214) 692-3112
Southern Methodist University, Elec. Eng. Dept., Dallas, TX 75275-0335

STEARNS, Mr. Steve .... (415) 795-5331
Tech. for Comms. Intl., 34175 Ardenwood Blvd., Fremont, CA 94536-7705

STEIN, Dr. Seymour .... (617) 527-5551
SCPE Inc., 56 Great Meadow Road, Newton, MA 02159

STOUT, Dr. David .... (408) 473-4690
Ford Aerospace & Comm. Corp., WDL Division, 220 Henry Ford II Drive, P.O. Box 49041, San Jose, CA 95161-9041

TREICHLER, Dr. John R. .... (408) 749-1888
Applied Signal Technology, 160 Sobrante Way, Sunnyvale, CA 94086

WARD, Dr. Robert .... (213) 812-4124
TRW, Mail Code R12/2043, One Space Park , Redondo Beach, CA 90278

WEBER, Dr. Charles .... (213) 743-2407
Communication Sciences Institute, University of Southern California, Department of Elec Eng., Los Angeles, CA 90089-0272

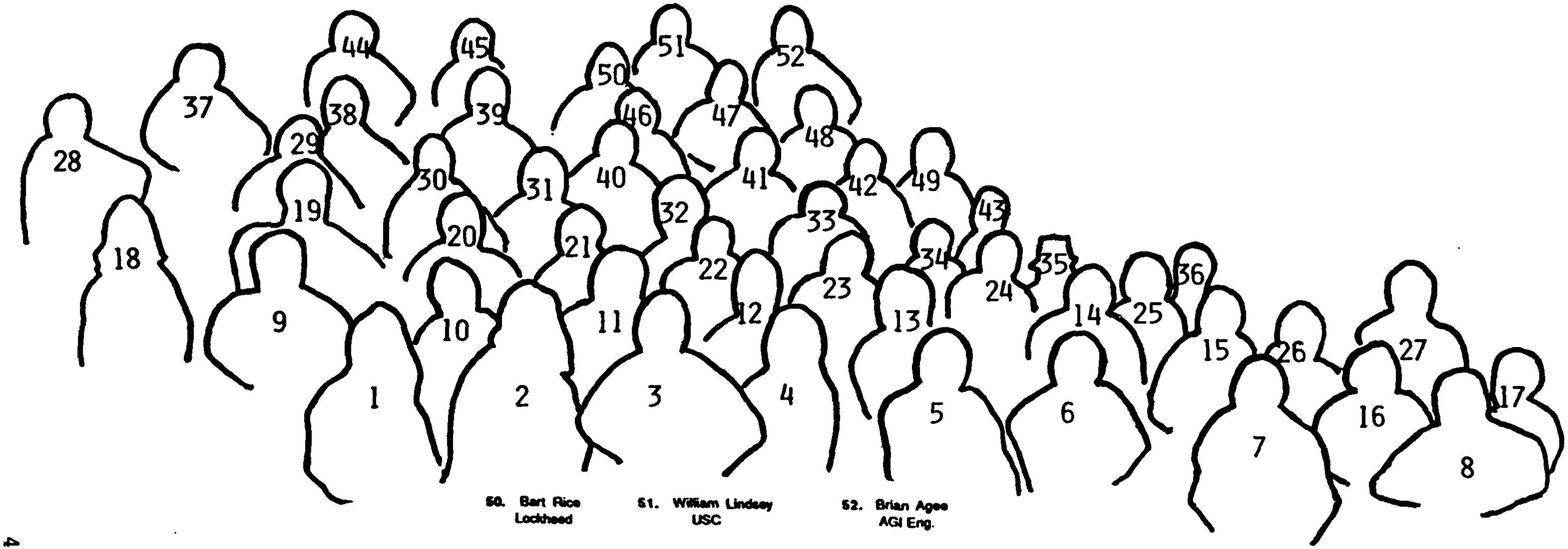
WELCH, Dr. Lloyd .... (213) 743-2699
Communication Sciences Institute, University of Southern California, Department of Elec Eng., Los Angeles, CA 90089-0272

WICKERT, Dr. Mark A. .... (719) 593-3500
University of Colorado 4-17400, College of Eng. & Applied Science, 1420 Austin Bluffs Parkway, Box 7150, Colorado Springs, CO 80907-7150

WILSON, Dr. Kenneth E. .... (303) 971-7545
Martin Marietta Corporation, P.O. Box 3912, Littleton, CO 80161

WOOD, Dr Sally L. .... (408) 749-1888
Dept. of Elec. Eng. & Computer Science, Santa Clara University, Santa Clara, CA 95053

ZHANG, Dr. Zhen .... 213) 743-3910
Communication Sciences Institute, University of Southern California, Department of Elec Eng., Los Angeles, CA 90089-0272



50. Bart Rice *Lockheed* 51. William Lindsey *USC* 52. Brian Agee *AGI Eng.*

44. Gaylord Huth *Axiomatix* 45. Steve Ames *Ford Aerosp.* 46. James Gault *ARO* 47. Andreas Polydoros *USC* 48. Raymond Pickholtz *George W. Univ.* 49. Seymour Stein *SCPE*

37. Ken Wilson *Martin Marietta* 38. Michael Pursley *U of Illinois* 39. William Sander *ARO* 40. William Gardner *UC Davis* 41. John Cioffi *Stanford U* 42. Samir Soliman *SMU* 43. Paul Sass *ARO*

28. Phillip Bello *MITRE* 29. Ben Friedlander *Sig. Proc. Tech.* 30. John Treicher *Appl. Sig. Tech.* 31. T.C. Huang *Aerospace Corpn.* 32. Mark Wickert *U of Colorado* 33. Allan Schneider *Cybercom* 34. Dennis Hall *TRW* 35. Stephen Rhodes *ARO* 36. Lloyd Welch *USC*

18. Milly Montenegro *USC* 19. Michael Rude *USC* 20. Robert Leyendecker *ARO* 21. Tom Dizer *ARO* 22. Peter Pawlowski *TRW* 23. Roger Peterson *Motorola* 24. Vedat Eyuboglu *Codex* 25. John Proakis *Northeastern U.* 26. David Stout *Ford. Aerosp.* 27. Zhen Zhang *USC*

9. Charles Bradshaw *MITRE Corpn.* 10.. Robert Ward *TRW* 11. Stuart Linsky *TRW* 12. Dean Kawaguchi *TRW* 13. Marvin Simon *JPL* 14. Ted Compton *Ohio State* 15. Allen Levesque *GTE* 16. Paul Crepeau *NRL* 17. P. Vijay Kumar *USC*

1. Sally Wood *UC, Santa Clara* 2. Cathy Cassells *USC* 3. Steve Stearns *TCI* 4. Monisha Ghosh *USC* 5. David Rollins *USC* 6. HermanBustamante *Stanford Telecomm.* 7. Robert Pelle *USC* 8. Robert Scholtz *USC*